

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2004 年 1 月 15 日 (15.01.2004)

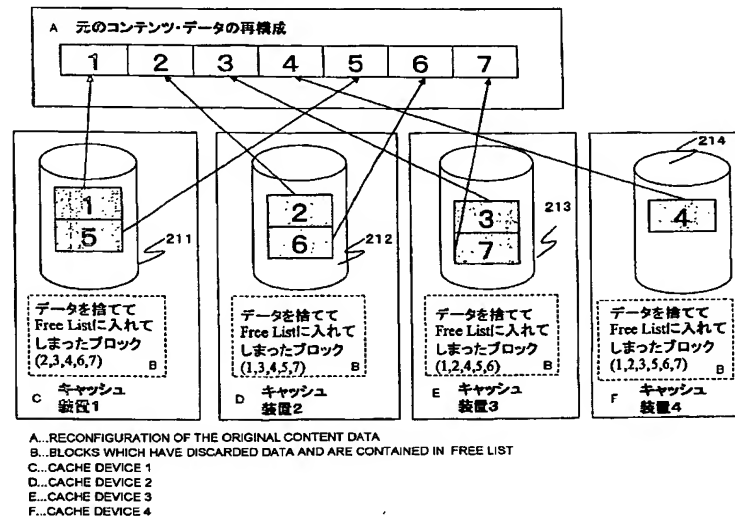
PCT

(10) 国際公開番号
WO 2004/006101 A1

- (51) 国際特許分類⁷: G06F 12/00, 13/00 (MIYAUCHI, Tutomu) [JP/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 ソニー株式会社内 Tokyo (JP).
- (21) 国際出願番号: PCT/JP2003/007971
- (22) 国際出願日: 2003 年 6 月 24 日 (24.06.2003)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願2002-195654 2002 年 7 月 4 日 (04.07.2002) JP
- (71) 出願人 (米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてののみ): 宮内 敦
- (74) 代理人: 宮田 正昭, 外 (MIYATA, Masaaki et al.); 〒104-0041 東京都中央区新富一丁目1番7号 銀座ティークエビル 澤田・宮田・山田特許事務所 Tokyo (JP).
- (81) 指定国 (国内): CN, KR, US.
- 添付公開書類:
— 国際調査報告書
- 2 文字コード及び他の略語については、定期発行される各 PCT ガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

(54) Title: CACHE DEVICE, CACHE DATA MANAGEMENT METHOD, AND COMPUTER PROGRAM

(54) 発明の名称: キャッシュ装置、およびキャッシュデータ管理方法、並びにコンピュータ・プログラム



(57) Abstract: A cache device and a cache data management method capable of effectively using a storage area and improving the hit ratio. In a cache replacement processing in the cache device connected to a network, a list of deletion-pending blocks is maintained and different data blocks are set to be deletion-pending for each cache device in a cache group so that data management is performed by applying the deletion-pending list. With this configuration, as compared to the cache replacement management using a single cache device, it is possible to effectively utilize the storage area. Moreover, in response to a data request from a terminal, it is possible to collect data blocks stored in a plurality of cache devices and transmit them, thereby reducing the network traffic and improving the hit ratio of the cache device.

(57) 要約: 効率的な記憶領域の利用、ヒット率の向上を実現するキャッシュ装置、キャッシュデータ管理方法を提供する。ネットワーク接続されたキャッシュ装置におけるキャッシュ・リフレッシュメント処理において、削除保留ブロックに関するリストを保持して、キャッシュグループ内のキャッシュ装置ごとに各々削除保留されるデータ・ブロックが異なるように設定して、削除保留リストを適用したデータ管理を行なう。本構成によれば、キャッシュ装置単独

[続葉有]

BEST AVAILABLE COPY

WO 2004/006101 A1



のキャッシュ・リプレースメント管理に比べて効率的な記憶領域の利用が図れ、また、端末からのデータ要求に応じて複数のキャッシュ装置に記憶したデータ・ブロックを集めて送信することが可能となり、ネットワークトラフィックの軽減、キャッシュ装置のヒット率の向上が実現される。

明 細 書

キャッシュ装置、およびキャッシュデータ管理方法、並びにコンピュータ・プログラム

5

技術分野

本発明は、キャッシュ装置、およびキャッシュデータ管理方法、並びにコンピュータ・プログラムに関する。さらに詳細には、ネットワークを通じて送受信されるデータ・キャッシュ処理において、キャッシュ装置内のデータ格納領域の効率的利用、およびヒット率向上を実現したキャッシュ装置、およびキャッシュデータ管理方法、並びにコンピュータ・プログラムに関する。

10

背景技術

15

近年、インターネット等の通信ネットワークを介した画像、音声データ、各種プログラム等のデータ転送が盛んに行われている。ネットワークを通じてやりとりされるコンテンツの種類は、テキスト・静止画などから、動画・音声などのマルチメディアコンテンツへと移行が進み、コンテンツの大容量化、多種類化が著しく進んでいる。

20

また、近年は、ブロードバンドによる大容量コンテンツの配信が急速に広まり、これに応じて、ネットワークを介して流通するコンテンツ数が飛躍的に増加している。このようにコンテンツの種類が増加し、大容量コンテンツが増大すると、データ通信量の増大、すなわちネットワークのデータトラフィックの増大という問題が発生する。

25

ネットワークのデータトラフィックの増大という問題を解決する1つの手法にキャッシュシステムがある。キャッシュシステムは、例えばコンテンツ配信サイトからコンテンツ受信ユーザー側に至るネットワーク経路上にデータ蓄積手段としてのキャッシュを構成したシステムであり、ユーザー側から要求

されたデータ（コンテンツ）がキャッシュに蓄積されている場合は、コンテンツ配信サイトから改めてデータ（コンテンツ）を送信することなく、データをキャッシュから取り出してユーザーに提供可能としたものである。

- 例えば、ユーザーがネットワークを介してコンテンツを取得し、ネットワーク上の一時的な記憶装置（キャッシュ）にコンテンツが格納された場合、キャッシュは一定の決まったルールでコンテンツ格納、削除処理を行なう。例えば
- 5 キャッシュの記憶容量の越えない範囲で最新のコンテンツが入力される毎に、アクセスの頻度が古いコンテンツを書き換える処理を実行する。通常、このキャッシュにおけるコンテンツ管理ルールはコンテンツの配信側やユーザー側の直接的な管理とは無関係に実行される。
- 10

このように、インターネット・キャッシュのデータ領域管理において、データが満杯になると使われなくなったデータ領域を検出し、その領域をキャッシュから削除後に新しいデータで置き換える処理を行なう。これをキャッシュ・リプレースメント（Cache Replacement）と呼ぶ。

- 15 従来のインターネット・キャッシュのキャッシュ・リプレースメント（Cache Replacement）は、インターネット等のネットワークに複数のキャッシュが存在する場合であっても、各々の個別のキャッシュ装置内で独自のデータ管理を実行するのが一般的である。最も一般的なキャッシュ・リプレースメントの技法としては、LRU（Least Recently Used）によるデータ管理方法がある。

- 20 LRUは、最近使用されていないデータ・ブロックを、LRUリストの後ろに回し、最近使われたデータ・ブロックを、LRUリストの先頭に設定するという管理方法であり、新しいデータをキャッシュするために古いデータを削除したい場合には、LRUリストの後ろにあるブロックから削除することで、キャッシュ内に新たなデータ格納領域を形成する方法である。

- 25 図1を参照してLRU（Least Recently Used）によるデータ管理方法について説明する。管理データとしては、利用頻度順にデータ・ブロックに対応するエントリを並べた（a）LRUリストと、キャッシュの空き領域ブロックを示す（b）フリーブロックリストを有する。新たにキャッシュされたデータは、（b）フリーブロックリストのエントリに示されるポインタによって特定され

るメモリ領域としてのデータ格納ブロックに格納され、格納ブロックに対して設定されるブロック情報を格納したエントリが、(a) LRUリストの先頭エントリ 101として設定される。

データの新規入力(キャッシュ)により、キャッシュのデータ格納領域が満杯になり、以前よりキャッシュされていたデータの中から、いずれかを削除する必要のあるときに、削除されるデータは、(a) LRUリストの末端にある最近使われていないデータ・ブロックと考えられるエントリ 102に対応するデータ・ブロックであり、削除されたデータ・ブロックをポインタ情報として示すエントリが(b)フリーブロックリストに設定されることで、所定のデータ格納ブロックが空き領域として設定される。

このようなキャッシュ用データ・ブロック管理の方法は、あくまで単独のキャッシュ装置に関する情報を元に処理されていた。また、近隣の複数のキャッシュ装置に対して要求元のキャッシュ装置の持っていないデータを要求するための手段としてICP(Internet Cache Protocol : RFC-2186, RFC-2187)が提案されているが、コンテンツファイルのリソース情報としてのURL(Uniform Resource Locator)を指定したURL単位でのデータ要求処理であることや、複数のキャッシュ間でどのデータを消去または残すべきかの細かい調整機能が無い。このため、たまたま各キャッシュ装置の事情で要求したものと同一データが他のキャッシュ装置に残っているときに、そのデータをサーバーからダウンロードしてくる代わりに他のキャッシュ装置から持ってくるという処理をしていただけである。

発明の開示

本発明は、上述したキャッシュ管理の現状に鑑みてなされたものであり、ネットワークに複数のキャッシュが接続された構成において、効率的なデータ格納領域の利用を実現したキャッシュ装置、およびキャッシュデータ管理方法、並びにコンピュータ・プログラムを提供することを目的とする。

例えば、ネットワーク接続されたキャッシュにおけるデータ領域管理方法において、データの削除を一時保留する「削除保留リスト」を使って管理することや、キャッシュグループという概念を導入することで、ビデオコンテンツのような大きなサイズのコンテンツをキャッシュするときに好適なデータ管理方法を実現する。

5 キャッシュグループとは、ネットワークに接続したキャッシュ装置のいくつかを組にしたものであり、このグループ内の複数のディスク間で同じコンテンツデータが管理されているものとする。このときに、あるデータ単位ごとにブロック化して各々のブロック単位で削除保留リストを使って管理するときに、
10 ディスク別に削除保留すべき部分が異なるように処理することにより、キャッシュ・リプレースメントによって、ある程度のデータ削除処理が進んだ段階においても、キャッシュグループ内の近隣のキャッシュディスクからデータを集めることによって、元のコンテンツデータを迅速に再生することができる。短いコンテンツデータであれば、コンテンツ別に削除保留にすべきものをディスク別に異なるように設定することで、同様の管理が可能である。

本発明の第1の側面は、

ネットワークを介して受信したデータをキャッシュデータとして格納し、端末からのデータ要求に応じてキャッシュデータを取得して端末に送信するキャッシュ装置であり、
20

自キャッシュ装置およびネットワークに接続された他のキャッシュ装置によって構成され、コンテンツの共同管理を実行するキャッシュグループに関するグループ設定情報を格納したキャッシュグループ・テーブルと、

前記キャッシュグループ・テーブルの格納情報に基づいて、コンテンツを構成する複数のデータ・ブロックから、自キャッシュ装置における記憶保持データとしての保留データ・ブロックを決定し、決定した保留データ・ブロックの対応情報からなるエントリによって構成される削除保留リストに従って保留データ・ブロックの管理を実行する制御手段と、
25

を有することを特徴とするキャッシュ装置にある。

さらに、本発明のキャッシュ装置の一実施態様において、前記キャッシュグループ・テーブルは、キャッシュグループによって共同管理するコンテンツの識別子と、キャッシュグループを構成するキャッシュ装置数：Nと、各キャッシュ装置に対応して設定されるグループ・メンバー番号：mとを有し、前記制御手段は、共同管理するコンテンツを構成する複数のデータ・ブロックから、自キャッシュ装置における記憶保持対象とする保留データ・ブロックの決定処理を、以下の処理、すなわち判定対象データ・ブロックが構成要素となるコンテンツの先頭データ・ブロックからのオフセット情報としてのブロックオフセット番号：Cに基づいて、

- 10 $C \bmod N$ を算出し、
 $C \bmod N = 0$ の場合は $V = N$ 、
 $C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、
 $V = m$ であるか否かを判定し、
 $V = m$ である場合に、該データ・ブロックを保留データ・ブロックとして決定する処理、
- 15 を実行する構成であることを特徴とする。

- さらに、本発明のキャッシュ装置の一実施態様において、前記制御手段は、記憶手段に格納したデータ・ブロックについての管理リストであるLRUリストからの排除対象エントリに対応するデータ・ブロックについて、前記キャッシュグループによる共同管理が可能か否かをキャッシュ装置間のメッセージ送受信により判定し、共同管理可能であると判定した場合に、前記キャッシュグループ・テーブルを生成する構成であることを特徴とする。
- 20

- さらに、本発明のキャッシュ装置の一実施態様において、前記制御手段は、
- 25 記憶手段に格納したデータ・ブロックについての管理リストであるLRUリストからの排除対象エントリに対応するデータ・ブロックについて、前記キャッシュグループ・テーブルの格納情報に基づいて、コンテンツを構成する複数のデータ・ブロック中から、自キャッシュ装置における記憶保持データとしての保留データ・ブロックを決定し、決定した保留データ・ブロック以外のデータ・

ブロックについては、削除の許容されるデータ・ブロック管理リストとしてのフリーブロックリストにより管理する構成であることを特徴とする。

さらに、本発明のキャッシュ装置の一実施態様において、前記削除保留リストは、データ・ブロックの保存優先度に応じた複数のリストによって構成され、
5 前記制御手段は、データ・ブロックの保存優先度を判定し、該判定結果に基づいて選択した削除保留リストに対して前記管理データ・ブロックに対応するエントリの設定を行なう構成であることを特徴とする。

さらに、本発明のキャッシュ装置の一実施態様において、前記制御手段は、端末からのデータ・ブロック取得要求が、前記キャッシュグループによる共同
10 管理対象のコンテンツに含まれるデータ・ブロック取得要求である場合に、取得要求対象のデータ・ブロックが前記キャッシュグループを構成するいずれのキャッシュ装置の保持対象データ・ブロックであるかを判定し、該判定結果に基づいてデータ・ブロックを自キャッシュ装置または、前記キャッシュグループを構成する他のキャッシュ装置から取得する処理を実行する構成であることを特徴とする。
15

さらに、本発明のキャッシュ装置の一実施態様において、前記キャッシュグループ・テーブルは、キャッシュグループによって共同管理するコンテンツの識別子と、キャッシュグループを構成するキャッシュ装置数： N と、各キャッシュ装置に対応して設定されるグループ・メンバー番号： m とを有し、前記制
20 御手段は、取得要求対象のデータ・ブロックが構成要素となるコンテンツの先頭データ・ブロックからのオフセット情報としてのブロックオフセット番号： C に基づいて、

$C \bmod N$ を算出し、

$C \bmod N = 0$ の場合は $V = N$ 、

25 $C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、

$V = m$ を算出し、

算出した m をグループ・メンバー番号として有するキャッシュ装置からデータ・ブロックの取得処理を実行する構成であることを特徴とする。

さらに、本発明の第2の側面は、

ネットワークを介して受信したデータをキャッシュデータとして格納し、端末からのデータ要求に応じてキャッシュデータを取得して端末に送信するキャッシュ装置におけるキャッシュデータ管理方法であり、

- 5 自キャッシュ装置およびネットワークに接続された他のキャッシュ装置によって構成され、コンテンツの共同管理を実行するキャッシュグループに関するグループ設定情報を格納したキャッシュグループ・テーブルの格納情報に基づいて、コンテンツを構成する複数のデータ・ブロックから、自キャッシュ装置における記憶保持データとしての保留データ・ブロックを決定する保留データ・ブロック決定ステップと、
- 10 決定した保留データ・ブロックの対応情報からなるエントリによって構成される削除保留リストに従って保留データ・ブロックの管理を実行する管理ステップと、

を有することを特徴とするキャッシュデータ管理方法にある。

- さらに、本発明のキャッシュデータ管理方法の一実施態様において、前記キャッシュグループ・テーブルは、キャッシュグループによって共同管理するコンテンツの識別子と、キャッシュグループを構成するキャッシュ装置数： N と、各キャッシュ装置に対応して設定されるグループ・メンバー番号： m とを有し、前記保留データ・ブロック決定ステップは、共同管理するコンテンツを構成する複数のデータ・ブロックから、自キャッシュ装置における記憶保持対象とする保留データ・ブロックの決定処理を、以下の処理、すなわち、
- 15
- 20

判定対象データ・ブロックが構成要素となるコンテンツの先頭データ・ブロックからのオフセット情報としてのブロックオフセット番号： C に基づいて、

$C \bmod N$ を算出し、

$C \bmod N = 0$ の場合は $V = N$ 、

- 25 $C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、

$V = m$ であるか否かを判定し、

$V = m$ である場合に、該データ・ブロックを保留データ・ブロックとして決定する処理を実行することを特徴とする。

さらに、本発明のキャッシュデータ管理方法の一実施態様において、前記キ

キャッシュデータ管理方法は、さらに、記憶手段に格納したデータ・ブロックについての管理リストであるLRUリストからの排除対象エントリに対応するデータ・ブロックについて、前記キャッシュグループによる共同管理が可能か否かをキャッシュ装置間のメッセージ送受信により判定し、共同管理可能であると判定した場合に、前記キャッシュグループ・テーブルを生成する処理を実行することを特徴とする。

さらに、本発明のキャッシュデータ管理方法の一実施態様において、前記キャッシュデータ管理方法は、さらに、記憶手段に格納したデータ・ブロックについての管理リストであるLRUリストからの排除対象エントリに対応するデータ・ブロックについて、前記キャッシュグループ・テーブルの格納情報に基づいて、コンテンツを構成する複数のデータ・ブロック中から、自キャッシュ装置における記憶保持データとしての保留データ・ブロックを決定し、決定した保留データ・ブロック以外のデータ・ブロックについては、削除の許容されるデータ・ブロック管理リストとしてのフリーブロックリストにより管理するステップを含むことを特徴とする。

さらに、本発明のキャッシュデータ管理方法の一実施態様において、前記削除保留リストは、データ・ブロックの保存優先度に応じた複数のリストによって構成され、前記キャッシュデータ管理方法は、さらに、データ・ブロックの保存優先度を判定し、該判定結果に基づいて選択した削除保留リストに対して前記管理データ・ブロックに対応するエントリを設定を行なうステップを含むことを特徴とする。

さらに、本発明のキャッシュデータ管理方法の一実施態様において、前記キャッシュデータ管理方法は、さらに、端末からのデータ・ブロック取得要求が、前記キャッシュグループによる共同管理対象のコンテンツに含まれるデータ・ブロック取得要求である場合に、取得要求対象のデータ・ブロックが前記キャッシュグループを構成するいずれのキャッシュ装置の保持対象データ・ブロックであるかを判定し、該判定結果に基づいてデータ・ブロックを自キャッシュ装置または、前記キャッシュグループを構成する他のキャッシュ装置から取得するデータ・ブロック取得処理を実行するステップを含むことを特徴とする。

る。

さらに、本発明のキャッシュデータ管理方法の一実施態様において、前記キャッシュグループ・テーブルは、キャッシュグループによって共同管理するコンテンツの識別子と、キャッシュグループを構成するキャッシュ装置数： N と、
5 各キャッシュ装置に対応して設定されるグループ・メンバー番号： m とを有し、前記データ・ブロック取得処理は、取得要求対象のデータ・ブロックが構成要素となるコンテンツの先頭データ・ブロックからのオフセット情報としてのブロックオフセット番号： C に基づいて、

$C \bmod N$ を算出し、

10 $C \bmod N = 0$ の場合は $V = N$ 、

$C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、

$V = m$ を算出し、

算出した m をグループ・メンバー番号として有するキャッシュ装置からデータ・ブロックを取得する処理として実行することを特徴とする。

15 さらに、本発明の第3の側面は、

ネットワークを介して受信したデータをキャッシュデータとして格納し、端末からのデータ要求に応じてキャッシュデータを取得して端末に送信するキャッシュ装置におけるキャッシュデータ管理処理を実行するコンピュータ・プログラムであり、

20 自キャッシュ装置およびネットワークに接続された他のキャッシュ装置によって構成され、コンテンツの共同管理を実行するキャッシュグループに関するグループ設定情報を格納したキャッシュグループ・テーブルの格納情報に基づいて、コンテンツを構成する複数のデータ・ブロックから、自キャッシュ装置における記憶保持データとしての保留データ・ブロックを決定する保留データ・ブロック決定ステップと、
25

決定した保留データ・ブロックの対応情報からなるエントリによって構成される削除保留リストに従って保留データ・ブロックの管理を実行する管理ステップと、

を有することを特徴とするコンピュータ・プログラムにある。

本発明の構成によれば、ネットワーク接続されたキャッシュ装置におけるキャッシュ・リプレースメント処理において、削除保留ブロックに関するリストを保持して、キャッシュグループ内のキャッシュ装置ごとに各々削除保留されるデータ・ブロックが異なるようにデータ管理を行なう構成としたので、各キャッシュ装置が単独でキャッシュ・リプレースメントの管理をしている場合に比べて効率的な記憶領域の利用が図れ、また、端末からのデータ要求に応じて複数のキャッシュ装置に記憶したデータ・ブロックを集めて送信することが可能となり、ネットワークトラフィックの軽減、キャッシュ装置のヒット率の向上が実現される。

さらに、本発明の多段の削除保留リストを適用したキャッシュ装置構成によれば、保持重要度の高いデータが削除されずにキャッシュに残る可能性が高くなるため、さらに、キャッシュヒット率の向上を達成することが可能となる。

なお、本発明のコンピュータ・プログラムは、例えば、様々なプログラム・コードを実行可能な汎用コンピュータ・システムに対して、コンピュータ可読な形式で提供する記憶媒体、通信媒体、例えば、CDやFD、MOなどの記録媒体、あるいは、ネットワークなどの通信媒体によって提供可能なコンピュータ・プログラムである。このようなプログラムをコンピュータ可読な形式で提供することにより、コンピュータ・システム上でプログラムに応じた処理が実現される。

本発明のさらに他の目的、特徴や利点は、後述する本発明の実施例や添付する図面に基づくより詳細な説明によって明らかになるであろう。なお、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

図面の簡単な説明

図1は、従来のLRUリストおよびフリーブロックリストによるキャッシュ

データの管理構成について説明する図である。

図 2 は、本発明のキャッシュ装置が適用可能なネットワーク構成について説明する図である。

5 図 3 は、本発明のキャッシュ装置におけるデータ管理構成について説明する図である。

図 4 は、本発明のキャッシュ装置におけるデータ管理構成について説明する図である。

図 5 は、本発明のキャッシュ装置間における送受信メッセージのフォーマット例について説明する図である。

10 図 6 は、本発明のキャッシュ装置における近隣キャッシュ装置リストの構成データについて説明する図である。

図 7 は、本発明のキャッシュ装置間における送受信メッセージとしてのグループ設定情報メッセージのフォーマット例について説明する図である。

15 図 8 は、本発明のキャッシュ装置において格納するキャッシュグループ・テーブルのデータ例について説明する図である。

図 9 は、本発明のキャッシュ装置においてデータ管理に利用するリストについて説明する図である。

図 10 は、本発明のキャッシュ装置においてデータ管理に利用するLRUリストについて説明する図である。

20 図 11 は、本発明のキャッシュ装置においてデータ管理に利用する削除保留リストについて説明する図である。

図 12 は、本発明のキャッシュ装置においてデータ管理に利用するフリーブロックリストについて説明する図である。

25 図 13 は、本発明のキャッシュ装置におけるグループ形成、データ管理処理について説明するフロー図である。

図 14 は、本発明のキャッシュ装置におけるグループ形成、データ管理処理に際して送受信されるメッセージについて説明するシーケンス図である。

図 15 は、本発明のキャッシュ装置におけるデータ取得送信処理について説明するフロー図である。

図 1 6 は、本発明のキャッシュ装置におけるリダイレクト処理によるデータ取得送信処理について説明するフロー図である。

図 1 7 は、本発明のキャッシュ装置におけるデータ取得送信処理に際して送受信されるメッセージについて説明するシーケンス図である。

5 図 1 8 は、本発明のキャッシュ装置において多段構成の削除保留リストを持つ構成を説明する図である。

図 1 9 は、本発明のキャッシュ装置における多段構成の削除保留リスト構成について説明する図である。

10 図 2 0 は、本発明の多段構成の削除保留リストを持つキャッシュ装置におけるグループ形成、データ管理処理について説明するフロー図である。

図 2 1 は、本発明のキャッシュ装置におけるハード構成例を示す図である。

発明を実施するための最良の形態

15

以下、本発明のキャッシュ装置、およびキャッシュデータ管理方法について、図面を参照して、詳細に説明する。

[キャッシュ装置および処理概要]

20

図 2 は、本発明のキャッシュ装置を適用可能なネットワークシステムの全体構成例を示す図である。この図 2 に示す構成は、一般的なインターネット・キャッシュ装置、サーバー装置、端末装置から成るシステムと同様である。本発明のキャッシュ装置は、従来のネットワークシステム構成での適用が可能であり、ネットワーク内の複数のキャッシュ装置でグループ設定を行ない、グループ内においてキャッシュデータの管理を実行する。

25

各キャッシュ装置 1 (2 1 1) ~ キャッシュ装置 4 (2 1 4) は、お互いにネットワークで接続されていると共に、コンテンツサーバー 2 0 1 から配信されるデータ (コンテンツ) を一時的に保存する機能を持つ。キャッシュ装置 2 1 1 ~ 2 1 4 のネットワーク下流側には、コンテンツを受信する複数の端末装

置、例えばユーザーPC、携帯通信端末等の各種端末を接続できる構成となっている。

キャッシュ装置1(211)には、コンテンツ受信端末としての端末221a～c・・・、キャッシュ装置2(212)には、端末231a～c・・・、キャッシュ装置3(213)には、端末241a～b・・・、キャッシュ装置4(214)には、端末251a～c・・・がそれぞれ接続される。

ある端末からコンテンツデータの取得要求があると、コンテンツサーバー201からデータを読み出すか、中間にあるキャッシュ装置からデータを読み出す。端末からのコンテンツ要求は、例えば、コンテンツファイルのリソース情報としてのURL(Uniform Resource Locator)を指定したコンテンツ取得要求メッセージの送信によって行なわれる。

キャッシュ装置に端末から取得要求のあったコンテンツのURLに対応するデータが保存されている場合は、このキャッシュ装置からデータを読み出す。キャッシュからのデータ読み出しにより、コンテンツサーバー201からキャッシュ装置間におけるデータ転送が省略可能となり、ネットワーク上のデータトラフィックが緩和される。

本発明の構成では、ネットワークで接続された複数のキャッシュ装置、例えば、図2に示すキャッシュ装置1(211)～キャッシュ装置4(214)を「キャッシュグループ」としてグループ化し、これら複数のキャッシュ装置1(211)～キャッシュ装置4(214)間で、コンテンツの共同管理、すなわち複数のキャッシュ装置1(211)～キャッシュ装置4(214)でのコンテンツの分散保持を実行する。以下、本発明におけるキャッシュ管理の構成、および方法について説明する。

図2のキャッシュグループに属する4つのキャッシュ装置1(211)～キャッシュ装置4(214)に、例えば図3に示すように、同じコンテンツデータがキャッシュされているとする。コンテンツは、データ・ブロック1～7からなる。図3のように各キャッシュ装置1(211)～キャッシュ装置4(214)で同じデータ・ブロック1～7を保存しているものとする。これらのデータ・ブロック1～7は、それぞれのキャッシュ装置1(211)～キャッシ

キャッシュ装置 4 (214) がそれぞれの接続端末からの要求に応じて、コンテンツサーバー 201 から受信し、一時記憶しているデータ、すなわちキャッシュデータである。

もし、従来と同様、個々のキャッシュ装置が単独で、キャッシュ・リプレイスメント処理を実行すると、個々のキャッシュ装置は、新しく記憶手段に格納、すなわちキャッシュすべきデータを保存するために、端末からの取得要求頻度の低いデータ・ブロックを削除することになる。この削除したデータを再度提供する場合には、既にキャッシュ装置にデータが無いために、コンテンツサーバー 201 へデータを取りに行くことになる。

10 本発明のキャッシュ装置は、複数のキャッシュ装置 1 (211) ~ キャッシュ装置 4 (214) で構成するキャッシュグループで、データを共同管理する。

本発明のキャッシュデータ管理方法では、同じコンテンツデータを持つ複数のキャッシュ装置を使ってキャッシュをグループ化してキャッシュするデータを管理する。本発明におけるキャッシュ装置のキャッシュ・リプレイスメントは、キャッシュグループを構成する複数のキャッシュ装置 1 (211) ~ キャッシュ装置 4 (214) を単位として実行する。

15 キャッシュグループを構成するキャッシュ装置ごとに削除されずに残っているデータが異なるデータとなるように制御されてデータ削除を実行する。従って、キャッシュ装置におけるデータ格納領域の効率的利用が可能となる。さらに、端末からのデータ要求に応答して、キャッシュグループの複数キャッシュ装置から各格納データを転送して端末に対する出力が可能となり、総合的にキャッシュ装置のヒット率を向上させることが可能となる。

20 具体例について、図を参照して説明する。図 3 のように、キャッシュグループを構成する複数のキャッシュ装置 1 (211) ~ キャッシュ装置 4 (214) の各々が、同じコンテンツデータ・ブロック 1 ~ 7 を格納していた場合、キャッシュグループを単位としたデータ管理により、キャッシュ装置ごとに削除されずに残っているデータが異なるデータとなるようにデータ削除、すなわちキャッシュ・リプレイスメントを実行する。

このキャッシュグループを単位としたキャッシュ・リプレイスメントの結果、

図4に示すように、キャッシュ装置1(211)は、データ・ブロック1, 5をキャッシュデータとして保持し、データ・ブロック2, 3, 4, 6, 7をフリーリスト(削除データ・ブロックに相当)に移動する。キャッシュ装置2(212)は、データ・ブロック2, 6をキャッシュデータとして保持し、データ・ブロック1, 3, 4, 5, 7をフリーリストに移動する。キャッシュ装置3(213)は、データ・ブロック3, 7をキャッシュデータとして保持し、データ・ブロック1, 2, 4, 5, 6をフリーリストに移動する。キャッシュ装置4(214)は、データ・ブロック4をキャッシュデータとして保持し、データ・ブロック1, 2, 3, 5, 6, 7をフリーリストに移動する。

- 10 このように、キャッシュグループを構成する複数のキャッシュ装置1(211)～キャッシュ装置4(214)全体としてキャッシュ・リプレースメントを行ない、コンテンツ共同管理を実行する。キャッシュグループ単位でのキャッシュ・リプレースメント処理の具体的処理態様について説明する。

15 あらかじめ決めた一定の時間間隔ごとに、各キャッシュ装置が独自に持っているLRUリスト(図1参照)から、キャッシュ・リプレースメントによって近々に消去される可能性のある順にブロックをスキャンし、その消去予定のブロックに対応するコンテンツを持つキャッシュ装置が近隣にあるかどうかを問い合わせ、さらに、問い合わせ先のキャッシュ装置が消去予定のコンテンツを持っている場合には、相互にキャッシュグループを組めるか否かを確認する。

- 20 このような問い合わせに適用する送受信メッセージは、例えばネットワーク接続されたキャッシュ間のメッセージ送受信プロトコルとして定義されたICP(Internet Cache Protocol)プロトコルが適用可能である。従来のICPの仕様中には、キャッシュグループに参加できるかどうかの確認機能についてまで規定されていないが、例えば、ICP中のオペコードの1つとして定義されているIP__OP__QUERYメッセージ中のOPTION Flag
- 25 を拡張して、キャッシュグループに参加できるかどうかの確認メッセージを送受信することが可能である。

ネットワーク接続されたキャッシュ間のメッセージ送受信プロトコルとして定義されたICP(Internet Cache Protocol)プロトコルのフォーマット

を図5(a)に、ICP中のオペコードの1つとして定義されているIP__OP__QUERYメッセージフォーマットを図5(b)に示す。ICPフォーマットは、図5(a)に示すように、オペコード、バージョン、メッセージ長、リクエストNo.、オプション、オプションデータ、送信側ホストアドレス、

5 ペイロードの各データフィールドを有する。

特定のコンテンツデータを持っているか、およびキャッシュグループに参加できるかを、近隣のキャッシュ装置に問い合わせる際には、コンテンツファイルのリソース情報としてのURL(Uniform Resource Locator)を指定して、近隣キャッシュ装置に問い合わせメッセージを送信する。この問い合わせメッセージの1形態として、ICP(Internet Cache Protocol)プロトコルのOPTION Flagの拡張を行ない、ICP__OP__QUERYで、コンテンツファイルのリソース情報としてのURLを指定するとともに、このコンテンツについてキャッシュグループのメンバーとして参加可能かの確認の意味を含むフラグ(例えば、0x20000000(ICP__FLAG__GROUP__REQ))を追加する。なお、この問い合わせメッセージは、この手法に限らず、他のフォーマット、プロトコルを使用することも可能である。

15 なお、このコンテンツ保持問い合わせは、問い合わせを行なうキャッシュ装置が、マルチキャスト(Multicast)を適用して、複数の近隣キャッシュ装置に一度に問い合わせることも可能である。

20 なお、本発明において、「近隣のキャッシュ装置」とは、ある1つのキャッシュ装置に対してキャッシュグループとして組む可能性のあるキャッシュ装置の候補であり、例えばデータ送受信処理を、あらかじめ定めた許容時間内に実行可能なキャッシュ装置として設定される。

25 「近隣のキャッシュ装置」の具体的な設定態様としては、例えば、キャッシュ装置の設定ファイルにキャッシュグループを組める可能性のある複数のキャッシュ装置を予め登録し、これらの登録キャッシュ装置にpingなどを使ってキャッシュ装置間のRTT(Round Trip Time)を測定し、RTTが許容時間範囲内にある複数のキャッシュ装置を「近隣のキャッシュ装

置」として設定することが可能である。

キャッシュグループの設定を実行するキャッシュ装置は、「近隣のキャッシュ装置」に関するリストを保持する。リストは、対象となるキャッシュ装置に関するデータとして、ホスト名、IPアドレス、ICPポート番号を保存したリストである。

キャッシュ装置リスト構成例を図6に示す。キャッシュ装置リストは、図6に示すように、近隣のキャッシュ装置として設定されたキャッシュ装置に関するデータとして、ホスト名、IPアドレス、ICPポート番号を保存したリストである。

コンテンツ保持の有無および、グループ参加可否についての問い合わせを近隣のキャッシュ装置に対して実行しようとするキャッシュ装置は、例えば、ICP_OP_QUERYで、あるコンテンツデータに対応するURLを指定して問い合わせメッセージをブロードキャスト、マルチキャスト等により送信する。問い合わせを実行したキャッシュ装置は、問い合わせメッセージに対する答えを待つことになる。

しかし、応答を無限に待つことはできないし、マルチキャストで問い合わせした場合は、どれだけ返事が戻るか不明であるため、あらかじめ決めた一定時間の範囲内において、ヒット(HIT)応答、すなわち問い合わせに係るURLに関連するコンテンツデータを持っているとの応答が何処からいくつ戻るかをカウントし内部のキャッシュグループ・テーブル(対象とするコンテンツデータの識別子(URL, コンテンツID等)、グループの構成メンバー数、グループID、キャッシュグループを構成するメンバーの(メンバー番号とIPアドレスリストの対データ))に記憶する。

応答メッセージは、問い合わせに係るURLに関連するコンテンツデータを持っている旨の応答として、例えばICPメッセージのICP_OP_HIT、また、コンテンツデータを持っていない旨の応答として、ICP_OP_MISSを用いた応答メッセージが適用可能である。

上述の問い合わせメッセージを受信したキャッシュ装置は、問い合わせを送信してきたキャッシュ装置とのキャッシュグループを構成するメンバーにな

ることを承諾し、さらに、キャッシュグループ形成までの間、問い合わせのあったコンテンツに対応するデータ・ブロックを削除せずに保持できる能力を持つ場合に、H I T 応答を返す。

- 5 問い合わせメッセージを送信したキャッシュ装置と、H I T 応答を返した 1 以上のキャッシュ装置とによって、問い合わせメッセージの指定コンテンツに関するキャッシュグループを構成する。

- 10 問い合わせメッセージを送信し、応答メッセージを受信したキャッシュ装置が、1 以上の H I T 応答を受信してキャッシュグループを設定した場合は、キャッシュグループを構成するキャッシュ装置数 (N)、当該キャッシュ装置のキャッシュグループを構成するキャッシュ装置 (メンバー) 識別子としてのグループ・メンバー番号 (m)、キャッシュグループでの処理対象コンテンツデータ識別子 (C I D) などのグループ設定情報を、キャッシュグループを構成する各キャッシュ装置に伝える。

- 15 なお、キャッシュ装置 (メンバー) 識別子としてのグループ・メンバー番号 (m) は、キャッシュ装置の I P アドレスに対応して設定され、キャッシュグループを構成する全てのキャッシュ装置 (メンバー) は、キャッシュグループを構成する他の全てのキャッシュ装置 (メンバー) 識別子としてのグループ・メンバー番号 (m) と I P アドレスの対応情報を取得する。

- 20 グループ設定情報の伝達には、上述の問い合わせメッセージと同様、I C P が適用可能である。例えば、I C P において、情報通知のための I C P _ O P _ N O T I F Y のようなオペコードを独自に追加し、ペイロード部分に図 7 のようなキャッシュ装置 (メンバー) のグループ・メンバー番号と、I P アドレスデータを格納してキャッシュグループを構成する各キャッシュ装置に送信し、通知することになる。なお、図 7 に示す例は、I P アドレスとして I P v 25 4 を適用した例である。

図 7 に示すように、ペイロード部には、キャッシュグループを構成する全て (N) のキャッシュ装置 (メンバー) のグループ・メンバー番号 (キャッシュ・メンバー # 1 ~ # N) と I P アドレス情報が格納される。これらの情報を受領することにより、キャッシュグループを構成する全てのキャッシュ装置 (メン

バー)は、キャッシュグループを構成する他の全てのキャッシュ装置(メンバー)のグループ・メンバー番号とIPアドレスの対応を知ることができる。

グループ設定情報を伝達したキャッシュ装置、およびグループ設定情報を伝達された、キャッシュグループに属するキャッシュ装置は、図8に示すような
5 グループ設定情報を格納したキャッシュグループ・テーブルを自己の管理情報としてメモリに格納する。

キャッシュグループ・テーブルには、上述のグループ設定情報に基づいて取得される情報、すなわち、設定キャッシュグループにおいて共同管理するコンテンツの識別子、キャッシュグループの構成キャッシュ装置数[N]、キャッシュグループの識別子(ID)、およびキャッシュグループのメンバーリスト
10 として、メンバー番号とIPアドレスの対応データとが格納される。

このようにして作成したキャッシュグループを構成する各キャッシュ装置は、単独のデータ(コンテンツ)管理ではなく、グループ全体でのコンテンツデータの管理を行なう。以下、キャッシュグループ内の各キャッシュ装置にお
15 けるコンテンツ管理方法について説明する。

キャッシュグループ内の各キャッシュ装置は、図9に示すように、利用頻度順にデータ・ブロックを並べた(a)LRUリストと、キャッシュの空き領域ブロックをポインタ情報として持つ(c)フリーブロックリスト、さらに、(b)削除保留リストを有する。

削除保留リストとは、LRUリストから排除、例えば削除対象ブロックとなったデータ・ブロックの内、一時的に削除保留状態に設定したデータ・ブロックを管理するリストである。LRUリストから排除するエントリに対応するデータ・ブロックを削除対象として、フリーブロックリストで管理するか、あるいは削除を保留すべきブロックとして、削除保留リストで管理するかの判定は、
20 以下の処理手順によって行なう。

削除保留状態に設定するか否かの判定は、以下の処理手順について説明する。判定対象コンテンツの先頭からのブロックオフセット番号をC、キャッシュグループを構成するキャッシュ装置数をN、キャッシュグループにおける自己に設定されたグループ・メンバー番号をmとしたときに、

- (1) $C \bmod N$ を算出する。
- (2) $C \bmod N = 0$ の場合は $V = N$ 、
 $C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、
- (3) $V = m$ であれば、削除保留状態に設定する。

5 $V \neq m$ であれば、削除保留状態に設定せず削除する。

上述の処理を、キャッシュグループを構成するキャッシュ装置の各々で実行することにより、先に図4を参照して説明したように、キャッシュ装置間で、重複の無いデータ・ブロックの格納構成が実現される。

すなわち、 $m = 1$ であるキャッシュ装置1 (2 1 1) は、データ・ブロック
10 1, 5を削除保留状態として保持し、削除保留リストの管理ブロックとし、データ・ブロック2, 3, 4, 6, 7を削除対象としてフリーリストで管理する。
 $m = 2$ であるキャッシュ装置2 (2 1 2) は、データ・ブロック2, 6を削除保留状態として保持し、削除保留リストの管理ブロックとし、データ・ブロック1, 3, 4, 5, 7をフリーリストで管理する。 $m = 3$ であるキャッシュ装置
15 3 (2 1 3) は、データ・ブロック3, 7を削除保留状態として保持し、削除保留リストの管理ブロックとし、データ・ブロック1, 2, 4, 5, 6をフリーリストで管理する。 $m = 4$ であるキャッシュ装置4 (2 1 4) は、データ・ブロック4を削除保留状態として保持し、削除保留リストの管理ブロックとし、データ・ブロック1, 2, 3, 5, 6, 7をフリーリストで管理する構成が実
20 現される。

上述の処理の詳細について、キャッシュグループ内の各キャッシュ装置の有するLRUリスト、削除保留リスト、フリーブロックリストの内部構造を参照しながら説明する。

まず、LRUリストの構成について図10を参照して説明する。図10に示
25 すように、各キャッシュ装置は、自己のキャッシュデータを格納する記憶手段としてのデータ格納領域320、およびデータ格納領域に格納したデータに対するリストとしてのLRUリスト310を有する。

LRUリスト310は、各データ・ブロックに対応付けて生成される。ここでは、(a) ~ (p) のリストデータ (エントリ) が各データ・ブロックに対

5 応付けられて保持されている。LRUリストの内部構造は、図10に示すように、データ・ブロックへのポインタ311、コンテンツ識別子312、ブロックオフセット313、LRUリストの前リストデータ（エントリ）のポインタ314、LRUリストの後続リストデータ（エントリ）のポインタ315を有する。

データ・ブロックへのポインタ311は、各リストデータに対応する実コンテンツデータ・ブロックの、データ格納領域320における格納アドレスを示す。コンテンツ識別子312は、例えばURL、または各データ・ブロックの識別データによって構成されるコンテンツIDである。ブロックオフセット313は、各コンテンツデータ・ブロックのコンテンツ先頭からのオフセットを示す。

15 オフセットは、例えばデータ格納領域におけるアドレス差分を用いることも可能であるが、ここでは、ある大きさごとにデータを区切ったブロックという単位で管理することになると、先頭からのブロックオフセット番号Cの値が入る。(a)のリストに対して設定されるブロックオフセットは1、(b)のリストに対して設定されるブロックオフセットは2となり、(p)のリストに対して設定されるブロックオフセットは、コンテンツデータのブロックサイズがNブロック分あるとすると、Nとなる。このようにブロックオフセットは、コンテンツ毎の、コンテンツ先頭からの位置識別データとしてのブロックオフセット番号Cとして設定される。

20 LRUリストの前リストデータ（エントリ）のポインタ314、LRUリスト後リストデータ（エントリ）のポインタ315は、それぞれ、リストの前後のエントリに対するポインタを示すデータである。従って、リスト先頭のリストデータ（a）の前リストデータのポインタはNULL、リスト最後尾のリストデータ（p）の後リストデータのポインタもNULLとして設定される。

LRUリスト中の各エントリが持つデータ・ブロックへのポインタに対応するコンテンツの識別子および先頭からのブロックオフセット番号：Cは、このエントリ中のデータを読むことで簡単に知ることが出来る。

次に、削除保留リストの構成について、図11を参照して説明する。削除保

留リストとは、前述したように、LRUリストで削除対象ブロックとなったデータ・ブロックの内、一時的に削除保留状態に設定したデータ・ブロックを管理するリストである。

図11に示すように、キャッシュ装置の保有する削除保留リスト330は、
5 削除保留状態に設定したデータ・ブロックに対応付けて生成される。ここでは、
(a)～(p)のリストデータ(エントリ)が削除保留状態に設定したデータ・
データ・ブロックに対応付けられて保持されている。削除保留リストの内部構造は、図11に示すように、データ・ブロックへのポインタ331、コンテンツ識別子332、ブロックオフセット333、設定されたキャッシュグループ
10 のグループ識別子(GP)334、削除保留リストの前リストデータ(エントリ)のポインタ335、削除保留リストの後続リストデータ(エントリ)のポインタ336を有する。

キャッシュ装置の保有する削除保留リスト330の格納データは、先に説明したLRUリストの格納データにグループ識別子(GP)334を付加した構成となる。グループ識別子(GP)334は、キャッシュグループ作成時に設定されるキャッシュグループを識別するグループIDである。キャッシュグループを設定したキャッシュ装置によって設定され、キャッシュグループを構成する他のキャッシュ装置は、キャッシュグループを設定したキャッシュ装置から送信されるグループ設定情報により通知され、キャッシュグループ・テーブル
20 (図8参照)に記録されたIDである。

次に、フリーブロックリストの構成について、図12を参照して説明する。フリーブロックリストとは、前述したように、LRUリストで削除対象ブロックとなったデータ・ブロックの内、削除保留状態に設定されずに削除対象となったデータ・ブロック、および削除保留状態に一旦、設定されたものの、その後、削除対象となったデータ・ブロックに対応する管理リストである。
25

図12に示すように、キャッシュ装置の保有するフリーブロックリスト350は、データ格納領域320のデータ格納ブロック中、上書き、消去可能なデータ・ブロックへのポインタ351を持ち、フリーブロックリストの前リストデータ(エントリ)のポインタ352、フリーブロックリストの後続リストデ

ータ（エントリ）のポインタ 3 5 3 を有する。

次に、キャッシュグループを構成する各キャッシュ装置において、新規入力データが発生し、データ・ブロックを削除保留リスト、あるいはフリーブロックリストの管理データに移行する際の処理手順について説明する。

- 5 まず、キャッシュ装置は、LRUリストから処理対象とするデータ・ブロックに対応するエントリ（リストデータ）を抽出し、抽出したLRUリストのエントリからブロックオフセット番号：C（図 10：ブロックオフセット 3 1 3）を読み出す。

- 10 次に、キャッシュ装置は、ブロックオフセット番号：Cの値に基づいて、上述した処理を実行する。すなわち、

（1） $C \bmod N$ を算出する。

（2） $C \bmod N = 0$ の場合は $V = N$ 、

$C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ 、

として、

- 15 （3） $V = m$ であれば、削除保留状態に設定、すなわち、

対応するデータ・ブロックの管理リストデータを削除保留リストに移動する。このデータ・ブロックは、データ格納領域において消去されずに保持される。

（3'）一方、 $V \neq m$ であれば、削除保留状態に設定しない。すなわち、

- 20 対応するデータ・ブロックの管理リストデータをフリーブロックリストに移動する。このデータ・ブロックは、データ格納領域において消去、上書き可能ブロックとして設定されることになる。

- 25 $V = m$ である場合は、設定したキャッシュグループにおいて、自装置がそのデータ・ブロックを保留すべきであると判定し、LRUリストからの削除予定エントリを削除保留リストに移動し、コンテンツデータ・ブロックを削除することなく、メモリに残し、削除保留リストでの管理を実行する。また、 $V \neq m$ の場合は、設定したキャッシュグループにおいて、自装置がデータ・ブロックを保留する必要がない、すなわち、そのデータ・ブロックは、他のグループ・メンバーにおいて保持すべきデータ・ブロックであると判定し、LRUリストからの削除予定エントリをフリーブロックリストに移動して、対応データ・ブ

ロックを削除可能状態とするものである。なお、削除されたデータ・ブロックは、キャッシュグループの他のメンバーが保持しており、これを取得して端末に送信可能となる。この処理については後述する。

リストに対するエントリの移動は、図 9 に示す処理となる。例えばキャッシュ装置 1 において、新規データ・ブロックの書き込み要求があると、(c) フリーブロックリストの先頭エントリ 271 に設定されたデータ・ブロック・ポインタに対応するデータ格納領域のデータ記憶領域に新規データ・ブロックが書き込まれ、書き込まれたデータ・ブロックに関する情報は、図 10 の LRU

- 10 その結果、もし LRU リストの長さが許容リスト長を超えることになった場合は、LRU リストの末尾のエントリ 273 が、削除保留リストあるいは、フリーブロックリストに移動することになる。どちらに移動するかの判定は、前述の (1) ~ (3) に示す $C \bmod N$ の算出、V の値の判定に従う。

- 15 上述した判定処理の結果、LRU リストの末尾のエントリ 273 を削除保留リストに移動すると決定した場合は、LRU リストの末尾のエントリ 273 が削除保留リストの末尾のエントリ 274 として設定される。また、上述した判定処理の結果、フリーブロックリストに移動すると決定した場合は、LRU リストの末尾のエントリ 273 がフリーブロックリストの末尾のエントリ 275 として設定される。

- 20 また、削除保留リストに新たなエントリが加わり、削除保留リストが予め定められたリスト長を超えることとなった場合は、削除保留リストの先頭のエントリ 276 が、フリーブロックリストの末尾のエントリ 275 として設定される。

- 25 なお、フリーブロックリストに対応して確保すべきバッファ量が、予め定められており、これが、新規データの格納により不足することとなった場合は、LRU リストあるいは削除保留リストからのフリーブロックリストへの移動、すなわちデータ・ブロックの削除によるバッファ量確保処理が実行される。

フリーブロックリストへ移動されたエントリの指定するデータ・ブロック・ポインタの指定するメモリ領域は、新たな入力データの書き込みに適用され、

データ・ブロックの内容が捨てられることになる。

[キャッシュグループ形成およびデータ管理処理]

次に、上述したキャッシュグループにおけるグループ形成、データ管理の手
5 順について、フローチャートおよびシーケンス図を参照して説明する。

まず、キャッシュグループの生成処理、データ管理手順について、図 1 3 お
よび図 1 4 を参照して説明する。

キャッシュグループは、自己のキャッシュ装置に格納したコンテンツのデー
タ・ブロックをLRUリストから削除保留リストによる管理に移動して、グル
10 ープでの共同管理を行なう場合に生成される。図 1 3 のフローに従って、キャ
ッシュグループの生成処理手順について説明する。

まず、ネットワーク接続されたキャッシュ装置は、事前準備処理として、キ
ャッシュグループのメンバー候補の登録を行なう。これは、図 1 3 におけるス
テップ S 1 0 1 ~ S 1 0 4 のステップである。

15 ステップ S 1 0 1 において、キャッシュグループを組む可能性のあるキャッ
シュ装置情報を登録する。これは、例えばオペレータが選択した 1 以上のキャ
ッシュ装置のホスト名、IP アドレス、ICP ポート番号の登録情報である。

次に、これらの登録情報から実際にキャッシュグループを設定可能なキャッ
シュ装置を判定する。これは、ステップ S 1 0 1 において登録された各キャッ
20 シュ装置に ping などを使ってキャッシュ装置間の R T T (R o u n d T
r i p T i m e) を測定し、R T T が許容時間範囲内にある複数のキャッ
シュ装置を選択するものであり、ステップ S 1 0 2 の R T T 測定処理、ステップ
S 1 0 3 の判定処理によって、あらかじめ定めた閾値以下の R T T を持つキャ
ッシュ装置をステップ S 1 0 4 において、「近隣のキャッシュ装置」として設
25 定する。

キャッシュ装置は、「近隣のキャッシュ装置」として設定されたキャッシュ
装置リストを保持する。リストは、先に図 6 を参照して説明したように、近隣
のキャッシュ装置に関するホスト名、IP アドレス、ICP ポート番号を保存
したリストである。

図13のステップS102～S104のステップに対応するデータ通信シーケンスが図14の[処理P]に相当する。図14のシーケンス図においては、キャッシュ装置Aが処理主体となるキャッシュ装置であり、キャッシュ装置B、C、Dが図13のステップS101において登録されたキャッシュ装置に相当する。

図14の[処理P]がRTT測定処理に相当する。キャッシュ装置Aは、キャッシュ装置B、C、Dに対して登録された各キャッシュ装置B、C、Dにpingなどを使ってキャッシュ装置間のRTT(Round Trip Time)を測定し、RTTが許容時間範囲内にある複数のキャッシュ装置を選択する。

図13のフローに戻り、説明を続ける。図6に示す近隣のキャッシュ装置リストが生成されると、キャッシュ装置は、ステップS105において、所定時間毎に、キャッシュ・リプレースメントにより、自己の保有するLRUリストから削除すべきコンテンツデータ・ブロックが発生したか否かの判定を実行する。

LRUリストから削除すべきデータ・ブロックがあると判定すると、ステップS106に進み、LRUリスト中の削除予定のエントリを抽出し、ステップS107において、そのエントリに対応するコンテンツがキャッシュグループ・テーブルに登録済みのコンテンツであるかを判定する。

キャッシュグループ・テーブルは、先に、図8を参照して説明したように、設定キャッシュグループにおいて共同管理するコンテンツの識別子、キャッシュグループの構成キャッシュ装置数[N]、キャッシュグループの識別子(ID)、およびキャッシュグループのメンバーリストとして、メンバー番号とIPアドレスの対応データとが格納されたテーブルであり、キャッシュグループ・テーブルに登録されたコンテンツは、設定されたキャッシュグループにおいて共同管理が実行されていることを示す。

ステップS107において、LRUリストから削除予定のコンテンツがキャッシュグループ・テーブルに登録されている場合は、設定されたキャッシュグループにおいて共同管理が実行されているコンテンツであり、共同管理処理と

して、ステップS 1 1 5以下の処理を実行することになる。この処理の詳細については、後段で説明する。

ステップS 1 0 7において、LRUリストから削除予定のコンテンツがキャッシュグループ・テーブルに登録されていない場合は、ステップS 1 0 8以下
5 において、LRUリストから削除予定のコンテンツに対応する新たなキャッシュグループを生成する。

ステップS 1 0 8では、先のステップS 1 0 4で設定したRTTが閾値以下の近隣キャッシュ装置リスト（図6参照）に設定された各キャッシュ装置に対して、LRUリストから削除予定のコンテンツに関するキャッシュグループを
10 組むことが可能か否かの問い合わせメッセージを送信する。

この問い合わせメッセージは、LRUリストから削除予定のエントリに対応するコンテンツ保持の有無および、グループ参加可否についての問い合わせメッセージである。図14のシーケンス図では、[処理Q]のICP__OP__QUERYがこの問い合わせメッセージに相当する。

図13のステップS 1 0 9で、キャッシュ装置は、問い合わせメッセージに対する応答を待機し、ステップS 1 1 0で応答が、グループ構成可（HIT）であるか否かを判定し、ステップS 1 1 1において、グループ構成可（HIT）であるキャッシュ装置について、のデータをキャッシュグループ・テーブル（図8参照）に登録する。ステップS 1 1 2では、全てのグループ・メンバー候補
20 （図6の近隣キャッシュ装置リストに登録されたキャッシュ装置）に対する返事の確認が終了したかを判定して、キャッシュグループの仮設定を終える。

図14のシーケンス図では、[処理Q]のICP__OP__QUERY（問い合わせメッセージ）に対して、グループ構成可（HIT）のキャッシュ装置B、Cが、ICPメッセージのICP__OP__HITをキャッシュ装置Aに対する
25 応答メッセージとして送信し、グループ構成不可（MISS）のキャッシュ装置Dが、ICPメッセージのICP__OP__MISSをキャッシュ装置Aに対する応答メッセージとして送信している。この場合、キャッシュ装置Aは、キャッシュ装置B、Cと、キャッシュグループを構成する。

図13の処理フローのステップS 1 1 3では、キャッシュグループ設定可能

な数のキャッシュ装置からのH I T応答が取得できたかを判定する。例えば1以上のキャッシュ装置からH I T応答が得られればキャッシュグループ設定可能と判定することができる。キャッシュグループ設定可能な数のキャッシュ装置のH I T応答がなかった場合は、ステップS 1 1 8に進み、キャッシュグループは設定せずにL R Uリストからの削除エントリをフリーブロックリストに移動する。この場合は、当エントリに対応するコンテンツのキャッシュグループでの共同管理は実行されない。

ステップS 1 1 3で、キャッシュグループ設定可能な数のキャッシュ装置からのH I T応答が取得できたと判定された場合は、ステップS 1 1 4において、H I T応答を返したキャッシュ装置であり、キャッシュグループ・テーブル(図8参照)にキャッシュグループのメンバーとして登録されたキャッシュ装置に対してグループ設定情報を送信する。

グループ設定情報は、キャッシュグループを構成するキャッシュ装置数(N)、当該キャッシュ装置のキャッシュグループを構成するキャッシュ装置(メンバー)識別子としてのグループ・メンバー番号(m)、キャッシュグループでの処理対象コンテンツデータ識別子などの情報である。グループ・メンバー番号(m)は、キャッシュ装置のI Pアドレスに対応して設定され、キャッシュグループを構成する全てのキャッシュ装置(メンバー)は、キャッシュグループを構成する他の全てのキャッシュ装置(メンバー)識別子としてのグループ・メンバー番号(m)とI Pアドレスの対応情報を取得する。

図14のシーケンス図において、[処理R]のI C P _ O P _ N O T I F Yが、グループ設定情報の送信に相当し、この送信データフォーマットは、例えば先に図7を参照して説明したフォーマットに従ったものとなる。グループ設定情報を伝達したキャッシュ装置、およびグループ設定情報を伝達された、キャッシュグループに属するキャッシュ装置は、図8に示すようなグループ設定情報格納したキャッシュグループ・テーブルを自己の管理情報としてメモリに格納する。

次にステップS 1 1 5以下が、設定されたキャッシュグループにおけるコンテンツの共同管理処理として実行されるL R Uリストからの移動処理である。

キャッシュ装置は、ステップS 1 1 5において、LRUリストからの削除予定エントリについて、ブロックオフセット番号：C、キャッシュグループを構成するキャッシュ装置数：N、キャッシュグループにおける自己に設定されたグループ・メンバー番号：mの各値に基づいて、

- 5 (1) $C \bmod N$ を算出し、
 (2) $C \bmod N = 0$ の場合は $V = N$ 、
 $C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、
 (3) $V = m$ であるか否かを判定する。

10 $V = m$ である場合は、設定したキャッシュグループにおいて、自装置がコンテンツを保留すべきであると判定し、ステップS 1 1 6において、LRUリストからの削除予定エントリを削除保留リストに移動する。すなわち、コンテンツデータ・ブロックを削除することなく、メモリに残し、削除保留リストでの管理を実行する。

15 $V = m$ でない場合は、設定したキャッシュグループにおいて、自装置がコンテンツを保留する必要がない、すなわち、他のグループ・メンバーにおいて保持すべきデータ・ブロックであると判定し、ステップS 1 1 8において、LRUリストからの削除予定エントリをフリーブロックリストに移動する。この場合は、対応データ・ブロックは、削除されることになる。削除された場合でも、キャッシュグループの他のメンバーが保持しており、これを取得して端末に送信可能となる。この処理については後述する。

20

さらに、ステップS 1 1 6におけるLRUリストから削除保留リストに対するエントリ移動により、削除保留リストのリスト長が許容値を超えることとなったとき(S 1 1 7でYes)は、ステップS 1 1 8において、削除保留リストのエントリをフリーブロックリストに移動する処理を実行する。

25

次に、図1 5乃至図1 7を参照して、キャッシュグループによるコンテンツ管理を実行しているキャッシュ装置に、端末からのコンテンツ要求が発生した場合の処理について説明する。

図1 5のステップS 2 0 1において、キャッシュ装置が端末からコンテンツ

取得要求を受信すると、キャッシュ装置は、ステップS 2 0 2において、要求コンテンツデータがキャッシュグループ・テーブル（図8参照）に登録されたコンテンツであるか否かを判定する。登録されていない場合は、グループ管理されていないコンテンツであり、通常の処理、すなわち自装置に格納されている場合は、コンテンツを取得して端末に送信し、自装置に格納されていない場合は、コンテンツをコンテンツサーバーから取得して端末に送信する。

要求コンテンツデータがキャッシュグループ・テーブル（図8参照）に登録されたコンテンツである場合は、グループ管理されているコンテンツであり、ステップS 2 0 3で、キャッシュグループ・テーブル（図8参照）からキャッシュグループの構成要素数Nと、自身のメンバー番号m 1を取得する。

ステップS 2 0 4において、端末からのコンテンツ取得要求中に含まれるコンテンツのオフセット情報に基づいて、ブロックオフセットを算出して、ブロックオフセット番号：Cを取得する。次に、ステップS 2 0 5において、キャッシュグループを構成するキャッシュ装置数：N、キャッシュグループにおける自己に設定されたグループ・メンバー番号：m 1の各値に基づいて、

(1) $C \bmod N$ を算出し、

(2) $C \bmod N = 0$ の場合は $V = N$ 、

$C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、

(3) $V = m 1$ であるか否かを判定する。

$V = m 1$ である場合は、要求されたコンテンツが自装置内に格納されていることになり、ステップS 2 0 6に進み、LRUリストまたは削除保留リストのエントリに基づいて要求コンテンツ名、ブロックオフセット番号から自装置内のデータ格納領域情報を取得し、ステップS 2 0 7において、データの読み出しを実行し、正常読み出し完了を確認（S 2 0 8）し、ステップS 2 0 9において、グループID（G p）= 0であるか否かを判定する。

グループID（G p）= 0は、設定されたキャッシュグループのメンバー離脱が発生して、キャッシュグループの維持を行わない際にキャッシュ装置が、キャッシュグループ・テーブルのグループID（G p）を書き換えることにより設定される。グループID（G p）が、0の場合は、キャッシュグループか

ら離脱したメンバーが多くなってグループを構成できないことを示すエントリであるため、ステップS 2 1 0において、エントリを削除保留リストではなくLRUリストの先頭に移動させる。

また、グループID(G_p) = 0でない場合は、ステップS 2 1 1において、
5 データ・ブロックがアクセスをされたことから、削除保留リストの先頭に該当のエントリを移す。

さらに、ステップS 2 1 2で、読み出しデータを要求端末に送信し、指定サイズデータの送信の完了を確認(S 2 1 3)し、未終了である場合は、ステップS 2 1 4において、ブロックオフセット番号： C の更新処理($C = C + 1$)
10 を実行して、ステップS 2 0 5以下の処理を繰り返し実行して端末に対するコンテンツ送信を実行する。

ステップS 2 0 5において、キャッシュグループを構成するキャッシュ装置数： N 、キャッシュグループにおける自己に設定されたグループ・メンバー番号： m_1 の各値に基づく、 $C \bmod N$ 算出、 $V = m_1$ であるか否かの判定
15 において、 $V = m_1$ でない場合は、要求されたコンテンツが自装置内に格納されていないことになり、ステップS 2 1 5に進み、算出した V の値(m_2 とする)に対応するグループ・メンバーをキャッシュグループ・テーブル(図8参照)から取得する。

ステップS 2 1 6において、算出した V の値(m_2 とする)に対応するグループ・メンバーのIPアドレスをキャッシュグループ・テーブル(図8参照)
20 から取得する。ステップS 2 1 7では、取得IPアドレスが0. 0. 0. 0か否かを判定する。IPアドレスが0. 0. 0. 0の場合は、キャッシュグループを離脱したメンバーのエントリを示す。この場合は、上流のキャッシュ装置またはコンテンツサーバーに要求をリダイレクトして発行(S 2 2 1)する。

25 ステップS 2 1 7の判定において、取得IPアドレスが0. 0. 0. 0でない場合は、キャッシュグループのメンバーのエントリを示す。ステップS 2 1 8に進み、コンテンツの読み出し要求(リダイレクト要求)をキャッシュグループのメンバー(m_2)に対応するIPアドレスに基づいて実行する。

リダイレクト要求を受信したキャッシュ装置における処理について、図16

を参照して説明する。これは、キャッシュグループを構成するキャッシュ装置が、他のグループ・メンバーからのコンテンツ読み出し要求（リダイレクト要求）に応答する処理として実行される処理である。

5 ステップS 3 0 1において、コンテンツ読み出し要求（リダイレクト要求）を受信すると、ステップS 3 0 2に進み、LRUリストまたは削除保留リストのエントリに基づいて要求コンテンツ名、ブロックオフセット番号から自装置内のデータ格納領域情報を取得し、ステップS 3 0 3において、データの有無、エラー状態を判定し、消去されているかエラーである場合は、ステップS 3 0 6において、コンテンツ読み出し要求（リダイレクト要求）を送信してきたキ
10 ッシュ装置に対してエラー提示コードを応答送信する。

 ステップS 3 0 3において、データの有無、エラー状態を判定し、データが存在し、読み出し可能であると判定されると、ステップS 3 0 4において、指定サイズのデータの読み出しを実行し、ステップS 3 0 5で、読み出しデータを要求キャッシュ装置に、正常終了提示コードとともに送信する。

15 図15の処理フローに戻り、説明を続ける。ステップS 2 1 9またはステップS 2 2 2において、リダイレクト要求先からの応答を受信すると、ステップS 2 2 0またはステップS 2 2 3において、正常読み出しがなされたか否かを判定し、Y e sの場合は、ステップS 2 0 9に進み、以下の処理、すなわち、リストエントリの移動、取得コンテンツの端末に対する送信等を実行する。な
20 お、ステップS 2 0 8、S 2 2 0、S 2 2 3において、データの正常読み出しに失敗した場合は、異常終了として処理を終了する。

 なお、異常終了として処理を終了する以前に、エラー応答受信後、リダイレクト先として上流のキャッシュ装置（サーバーへの経路途中にある、サーバーにより近いキャッシュ装置）がある場合は、上流の装置にリダイレクトし、無
25 い場合は、直接コンテンツサーバーに要求をリダイレクトして、目的とするデータを得る処理を実行してもよい。

 上記処理における各キャッシュ装置間のデータ通信シーケンスの概要を、図17を参照して説明する。なお、図17において、左端から、コンテンツ取得要求を送信する端末、

端末からのコンテンツ取得要求を受信するキャッシュ装置 A

キャッシュ装置 A とキャッシュグループを構成するキャッシュ装置 1

キャッシュ装置 A とキャッシュグループを構成するキャッシュ装置 2

キャッシュ装置 A の上流キャッシュ装置またはサーバー

5 である。

〔処理 S〕は、端末からの要求コンテンツデータを、キャッシュ装置 A が、
キャッシュグループ・メンバーとしてのキャッシュ装置 1, 2 からそれぞれ取
得して、端末に送信する処理シーケンスであり、URL を指定したコンテンツ
取得要求 (HTTP GET (URL)) を受信したキャッシュ装置が上述の
10 C mod N 算出、 $V = m1$ の算出に基づいて、コンテンツの対応データ・
ブロックを保持していると予測されるキャッシュ装置 1, 2 に対してコンテ
ンツ取得要求 (HTTP GET (URL) Range: bytes=a-a-bb) を送信し、コンテンツのデータ・ブロックを取得 (HTTP 206) し、
キャッシュ装置 A は、取得コンテンツを端末に送信 (HTTP OK) する処
15 理シーケンスである。

〔処理 T〕は、端末からの要求コンテンツデータを、キャッシュ装置 A が、
キャッシュグループ・メンバーとしてのキャッシュ装置 1, 2 からそれぞれ取
得できなかった場合の処理シーケンスであり、URL を指定したコンテンツ取
得要求 (HTTP GET (URL)) を受信したキャッシュ装置が上述の C
20 mod N 算出、 $V = m1$ の算出に基づいて、コンテンツの対応データ・ブ
ロックを保持していると予測されるキャッシュ装置 1 に対してコンテンツ取得
要求 (HTTP GET (URL) Range: bytes=a-a-bb) を送信し、コンテンツのデータ・ブロックが取得できない旨の応答 (HTTP
404 (Not Found) または、HTTP 410 (Gone)) を受信す
25 る。

さらに、キャッシュ装置 A は、キャッシュ装置 A の上流キャッシュ装置または
サーバーに、コンテンツ取得要求 (HTTP GET (URL) Range:
bytes=a-a-bb) を送信し、コンテンツのデータ・ブロックを取得 (H
TTP 206) し、キャッシュ装置 A は、取得コンテンツを端末に送信 (HT

TP OK及び取得したコンテンツデータの送信)をする。

なお、[処理T]のように、キャッシュグループの構成メンバーから予定のコンテンツが取得できなかった場合は、キャッシュグループの構成メンバーとして存続する意味がないので、グループ解消を行なう。この処理が[処理U]

5 である。

グループ解消処理は、キャッシュAから、グループ解消通知(ICP OP NOT I F Y)を、グループを構成する各キャッシュ装置1, 2に送信することによって行なわれる。キャッシュ装置A, 1, 2はその後、キャッシュグループ・テーブルから対応するグループ情報を削除する。

10 なお、上述した例では、HTTPプロトコルを使ったリダイレクト手順を書いているが、FTPやRTSPなどの別のプロトコルであってもかまわない。また、キャッシュ装置間の通信プロトコルについても、ICPだけに限定する必要はなく、一般的に使われる装置間の通信手段を拡張することでも実現可能である。また、キャッシュ間通信プロトコルも、ICPだけには限らない。

15 キャッシュグループを構成するグループ・メンバー m ($m=1, 2, \dots, N$: N はグループ構成メンバー数)がキャッシュグループから離脱することになるケースを以下に説明する。

キャッシュグループを組むと宣言したコンテンツに関連するデータ・ブロックの全体数を A としたときに、グループ・メンバーとして保持しておくべきデータ・ブロック ($V=m$ のデータ・ブロック)の総数 S は、 $A \bmod N$ が
20 零の場合は $V=N$ 、非零の場合は $V=(A \bmod N)$ として $m \leq V$ のとき
 $S=(A+N-1)/N$ 、 $m > V$ のとき $S=A/N$ と計算される。

この S の分のデータ・ブロックを永久に保持できればよいが、記憶資源は有限であるため、キャッシュ・リプレースメントが進むとグループ・メンバーとして保持しておくべきデータ・ブロックであったとしても、削除保留リストから削除されてしまうケースもある。このようにグループ・メンバーとして保持しておくべきデータ・ブロックも削除保留リストからも削除されてしまうような場合において、グループ・メンバーとして保持すべきデータ・ブロックにもかかわらず削除してしまったデータ・ブロックの数を D とすると、グループ

離脱に至る閾値 T (%) として、グループ・メンバーとして保持すべきデータ・ブロックの残存率： $(S - D) / S \times 100$ (%) を算出し、

データ・ブロックの残存率と閾値 T との関係が、

$(S - D) / S \times 100$ (%) $< T$ ($0 \leq T < 100$) となった場合に、

- 5 そのコンテンツに関しキャッシュグループを組んでいるキャッシュ・メンバーに対して、グループ離脱を宣言する構成が可能である。

キャッシュグループからの離脱を検知したときは、自分のキャッシュグループ・テーブルから該当するコンテンツに対応するエントリを削除すると共に、
図 17 の [処理 U] に示すように、これを知らせるフラグ (説明のための例として 0×10000000 (ICP__FLAG__GROUP__DESERTED)) を立てた ICP__OP__NOTIFY メッセージを、該当するコンテンツについてキャッシュグループを組んでいる全グループ・メンバーに送る。

さらに、各キャッシュ装置にあるキャッシュグループ・テーブルから、通知されたキャッシュ装置がキャッシュグループから離脱したことがわかるように、
15 対応する部分の IP アドレスを "0. 0. 0. 0" のように書き換える。

ただし、メンバー数などの他の情報は、離脱メンバーが無い場合とのアルゴリズムの互換を保つ目的もあり、そのままの情報として保持してもよい。もし、離脱グループ・メンバーが保持しているべきデータ・ブロックをアクセスした場合には、オリジナルデータを持っているサーバーまたは上流のキャッシュ
20 装置にアクセスすることになる。

キャッシュグループを構成するメンバー数が、離脱メンバーを除いた実質数で 1 となった場合、キャッシュグループを組めないことになるため、グループを完全に消去する意味で、残った 1 台のキャッシュ装置のキャッシュグループ・テーブルから該当するコンテンツに対応するエントリを削除し、削除保留
25 リストの対応するコンテンツに属するデータ・ブロックの "Gp" の項目を初期化 (ゼロクリア) する。

[多段構成の削除保留リスト]

次に、キャッシュ装置においてデータ管理リストとして生成する削除保留リ

ストを多段構成とした実施例について説明する。

上述した実施例において、キャッシュ装置は、図 9 を参照して説明したように、LRUリストと、キャッシュの空き領域ブロックを示すフリーブロックリスト、さらに、LRUリストで削除対象ブロックとなったデータ・ブロックの内、一時的に削除保留状態に設定したデータ・ブロックを管理するリストとしての削除保留リストを有しており、各リストは、先頭エントリからシーケンシャルに並べられたエントリを有する構成であった。

本実施例のキャッシュ装置は、図 18 に示すように、複数の削除保留リストを有する。キャッシュ装置は、利用頻度順にエントリを並べた (a) LRUリスト、キャッシュグループを構成したコンテンツのデータ・ブロックに対応して設定したエントリを格納した複数の (b 1) (b 2) 削除保留リスト、空きブロック領域をポインタ情報として有する (c) フリーブロックリストを有する。

複数の削除保留リスト (b 1) (b 2) は、コンテンツデータ・ブロックの削除保留優先度の高低に基づいて振り分けられてエントリが各リストに格納される。ここでは、2つの削除保留リスト (b 1) (b 2) のうち、(b 1) のリストに保存優先度の高いコンテンツに対応するエントリを格納し、(b 2) のリストに保存優先度の低いコンテンツに対応するエントリを格納する例を示す。本例では、2つの削除保留リストの設定例を示しているが、保存優先度順に 3 以上の削除保留リストを設定する構成としてもよい。

なお、データ・ブロックの重要性に応じた保存優先度の判定の態様は、様々な態様が可能である。例えば、新規に届いたコンテンツに対応するデータ・ブロックほど重要とする設定、あるいは、コンテンツデータの先頭部分やシーンチェンジ部分以降の特定区間だけは重要度の高いデータ・ブロックとして設定するなどの構成がある。なお、コンテンツデータの先頭部分やシーンチェンジ部分の検出は、既存のソフトウェアによる処理が可能である。

リストに対するエントリの移動について説明する。例えばキャッシュ装置において、新規データ・ブロックの書き込み要求があると、(c) フリーブロックリストの先頭エントリ 501 に設定されたデータ・ブロック・ポインタに対

応するデータ格納領域のデータ記憶領域に新規データ・ブロックが書き込まれ、書き込まれたデータ・ブロックに関する情報は、図10のLRUリストの先頭エントリ502に移される。

その結果、もしLRUリストの長さが許容リスト長を超えることになった場合は、LRUリストの末尾のエントリ503が、削除保留リスト(b1)または(b2)あるいは、フリーブロックリスト(c)に移動することになる。削除保留リストかフリーブロックリストのどちらに移動するかの判定は、前述の $C \bmod N$ の算出、Vの値によるエントリに対応するコンテンツデータ・ブロックの残留、削除判定に基づくことになる。

上述した判定処理の結果、エントリを削除保留リストに移動すると決定した場合は、保存の重要度に応じて、重要度の高いデータ・ブロックに対するエントリは、削除保留リスト(b1)の末尾のエントリ504として設定し、重要度の低いデータ・ブロックに対するエントリは、削除保留リスト(b2)の末尾のエントリ505として設定する。

また、削除保留リスト(b1)に新たなエントリが加わり、削除保留リスト(b1)が予め定められたリスト長を超えることとなった場合は、削除保留リスト(b1)の先頭のエントリ511が、削除保留リスト(b2)の末尾のエントリ505として設定される。

同様に、削除保留リスト(b2)に新たなエントリが加わり、削除保留リスト(b2)が予め定められたリスト長を超えることとなった場合は、削除保留リスト(b2)の先頭のエントリ512が、フリーブロックリストの末尾のエントリ521として設定される。

なお、フリーブロックリストに対応して確保すべきバッファ量が、予め定められており、これが、新規データの格納により不足することとなった場合は、LRUリストあるいは削除保留リスト(b2)からのフリーブロックリストへの移動、すなわちデータ・ブロックの削除によるバッファ量確保処理が実行される。

フリーブロックリストへ移動されたエントリの指定するデータ・ブロック・ポインタの指定するメモリ領域は、新たな入力データの書き込みに適用され、

データ・ブロックの内容が捨てられることになる。

多段構成とした場合の削除保留リストの構成について、図 19 を参照して説明する。削除保留リストとは、前述したように、LRUリストで削除対象ブロックとなったデータ・ブロックの内、一時的に削除保留状態に設定したデータ・ブロックを管理するリストである。

図 19 は、図 18 の削除保留リスト (b 1) または (b 2) のいずれかに対応する削除保留リストの構成を示している。図 19 のように、キャッシュ装置の保有する削除保留リスト 630 は、削除保留状態に設定したデータ・データ・ブロックに対応付けて生成される。ここでは、(a) ~ (p) のリストデータ (エントリ) が削除保留状態に設定したデータ・データ・ブロックに対応付けられて保持されている。削除保留リストの内部構造は、図 19 に示すように、データ・ブロックへのポインタ 631、コンテンツ識別子 632、ブロックオフセット 633、設定されたキャッシュグループのグループ識別子 (GP) 634、さらに、削除保留リストの重要度指数値 (Priority) 635、および削除保留リストの前リストデータ (エントリ) のポインタ 636、削除保留リストの後続リストデータ (エントリ) のポインタ 637 を有する。多段構成の削除保留リスト 630 の格納データには、削除保留リストの重要度指数値 (Priority) 635 が付加されることになる。

次に、多段構成の削除保留リストを持キャッシュ装置におけるキャッシュグループの生成処理、データ管理手順について、図 20 を参照して説明する。

図 20 において、ステップ S 501 乃至ステップ S 515 は、前述の図 13 を参照して説明した処理フローのステップ S 101 乃至ステップ S 115 と同様の処理であるので、説明を省略する。

図 20 のステップ S 515 において、LRUリストからの削除予定エントリについて、ブロックオフセット番号: C、キャッシュグループを構成するキャッシュ装置数: N、キャッシュグループにおける自己に設定されたグループ・メンバー番号: m の各値に基づいて、 $C \bmod N$ を算出し、さらに、 $C \bmod N = 0$ の場合は $V = N$ 、 $C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、 $V = m$ であるか否かを判定した結果、 $V \neq m$ である場合は、

ステップS 5 2 3に進みそのデータ・ブロックをフリーリストで管理する。V = mである場合は、そのデータ・ブロックを削除保留として、自装置で管理することになる。

- 5 この場合、ステップS 5 1 6において、削除保留対象データ・ブロックの保持優先度の判定を実行する。この保持優先度判定は、キャッシュ装置において予め定められた基準、例えば、前述したように、新規に届いたコンテンツに対応するデータ・ブロックを重要、あるいは、コンテンツデータの先頭部分やシーンチェンジ部分以降の特定区間を重要とした設定情報に基づいて判定する。
- 10 保存優先度の高いデータ・ブロックであると判定した場合は、ステップS 5 1 7に進み、保存優先度の高い削除保留リスト（図18の（b1））にエントリを移動する。保存優先度の低いデータ・ブロックであると判定した場合は、ステップS 5 2 0に進み、保存優先度の低い削除保留リスト（図18の（b2））にエントリを移動する。
- 15 さらに、ステップS 5 1 8において、保存優先度の高い削除保留リストから保存優先度の低い削除保留リストにエントリ移動が必要であるか否かを判定し、必要があると判定されると、ステップS 5 1 9において、保存優先度の高い削除保留リストから保存優先度の低い削除保留リストにエントリを移動する。
- 20 ステップS 5 2 1において、保存優先度の低い削除保留リストからフリーブロックリストにエントリ移動が必要であるか否かを判定し、必要があると判定されると、ステップS 5 2 2において、保存優先度の低い削除保留リストからフリーブロックリストにエントリを移動する。
- 25 以上の処理に従って、多段構成を持つ削除保留リストを適用したデータ管理が実行される。

[キャッシュ装置の構成]

次に、上述の実施例において説明したキャッシュ装置のハード構成例について説明する。

図 2 1 に、制御手段として C P U (Central Processing Unit) を備えたキャッシュ装置例を示す。なお、図 2 1 に示す構成例は 1 つの例であり、キャッシュ装置は、ここに示す構成以外のハードウェアを持つことも可能である。

図 2 1 に示す構成について説明する。C P U (Central Processing Unit) 7
5 0 1 は、各種実行プログラム、O S (Operating System) を実行するプロセッサである。R O M (Read-Only-Memory) 7 0 2 は、C P U 7 0 1 が実行するプログラム、あるいは演算パラメータとしての固定データを格納する。R A M (Random Access Memory) 7 0 3 は、C P U 7 0 1 の処理において実行される
10 プログラム、およびプログラム処理において適宜変化するパラメータの格納エリア、ワーク領域として使用される。また、近隣キャッシュ装置リスト、キャッシュグループ・テーブルの格納領域として使用される。

C P U 7 0 1, R O M 7 0 2, R A M 7 0 3 は C P U バス 7 1 1 によってデータ転送可能に接続され、バスブリッジ 7 0 4 を介して I / O バス 7 1 2 に接続される。I / O バス 7 1 2 には、ディスクインタフェース 7 0 5、ネットワーク
15 インタフェース 7 0 7, 7 0 8 が接続される。

ディスクインタフェース 7 0 5 は、ハードディスク、フレキシブルディスク、C D - R O M (Compact Disc Read Only Memory), M O (Magneto optical) ディスク, D V D (Digital Versatile Disc)、磁気ディスク等のディスク 7 0 6 に対するインタフェースであり、ディスク 7 0 6 には、キャッシュデータが格納
20 され、C P U 7 0 1 の制御の下に必要データが取得、格納される。

ネットワークインタフェース 7 0 7 は、外部ネットワークインタフェースであり、ネットワーク接続されたサーバー、他のキャッシュ装置とのデータ送受信のインタフェースである。ネットワークインタフェース 7 0 8 は、内部ネットワークインタフェースであり、ネットワーク接続された端末とのデータ送
25 受信のインタフェースである。

なお、明細書中において説明した各処理はハードウェア、またはソフトウェア、あるいは両者の複合構成によって実行することが可能である。一連の処理をソフトウェアにより実行させる場合には、そのソフトウェアを構成するプログラムが専用のハードウェアに組み込まれているコンピュータ、又は各種のプ

プログラムをインストールすることで各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどに、フレキシブルディスクやCD-ROM等のプログラム読み取り可能な記録媒体にプログラムを格納して提供してもよいし、インターネットなどの通信網を介してプログラムをダウンロードしてもよい。

具体的には、プログラムは記録媒体としてのハードディスクやROM (Read Only Memory)に予め記録しておくことができる。あるいは、プログラムはフレキシブルディスク、CD-ROM (Compact Disc Read Only Memory), MO (Magneto optical)ディスク, DVD (Digital Versatile Disc)、磁気ディスク、半導体メモリなどのリムーバブル記録媒体に、一時的あるいは永続的に格納（記録）しておくことができる。このようなリムーバブル記録媒体は、いわゆるパッケージソフトウェアとして提供することができる。

また、プログラムは、上述したようなリムーバブル記録媒体からコンピュータにインストールする他、ダウンロードサイトから、コンピュータに無線転送したり、LAN (Local Area Network)、インターネットといったネットワークを介して、コンピュータに有線で転送し、コンピュータでは、そのようにして転送されてくるプログラムを受信し、内蔵するハードディスク等の記録媒体にインストールすることができる。

なお、明細書に記載された各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されてもよい。また、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

以上、特定の実施例を参照しながら、本発明について詳解してきた。しかしながら、本発明の要旨を逸脱しない範囲で当業者が該実施例の修正や代用を成し得ることは自明である。すなわち、例示という形態で本発明を開示してきたのであり、限定的に解釈されるべきではない。本発明の要旨を判断するためには、冒頭に記載した特許請求の範囲の欄を参酌すべきである。

産業上の利用可能性

以上、説明してきたように、本発明の構成によれば、ネットワーク接続されたキャッシュ装置におけるキャッシュ・リプレースメント処理において、削除
5 保留ブロックに関するリストを保持して、キャッシュグループ内のキャッシュ
装置ごとに各々削除保留されるデータ・ブロックが異なるようにデータ管理を
行なう構成としたので、各キャッシュ装置が単独でキャッシュ・リプレースメ
ントの管理をしている場合に比べて効率的な記憶領域の利用が図れ、また、端
10 末からのデータ要求に応じて複数のキャッシュ装置に記憶したデータ・ブロッ
クを集めて送信することが可能となり、ネットワークトラフィックの軽減、キ
ャッシュ装置のヒット率の向上が実現される。

さらに、本発明の多段の削除保留リストを適用したキャッシュ装置構成によ
れば、保持重要度の高いデータが削除されずにキャッシュに残る可能性が高く
なるため、さらに、キャッシュヒット率の向上を達成することが可能となる。

請求の範囲

1. ネットワークを介して受信したデータをキャッシュデータとして格納
5 し、端末からのデータ要求に応じてキャッシュデータを取得して端末に送信するキャッシュ装置であり、

自キャッシュ装置およびネットワークに接続された他のキャッシュ装置によって構成され、コンテンツの共同管理を実行するキャッシュグループに関するグループ設定情報を格納したキャッシュグループ・テーブルと、

10 前記キャッシュグループ・テーブルの格納情報に基づいて、コンテンツを構成する複数のデータ・ブロックから、自キャッシュ装置における記憶保持データとしての保留データ・ブロックを決定し、決定した保留データ・ブロックの対応情報からなるエントリによって構成される削除保留リストに従って保留データ・ブロックの管理を実行する制御手段と、

15 を有することを特徴とするキャッシュ装置。

2. 前記キャッシュグループ・テーブルは、キャッシュグループによって共同管理するコンテンツの識別子と、キャッシュグループを構成するキャッシュ装置数：Nと、各キャッシュ装置に対応して設定されるグループ・メンバー
20 番号：mとを有し、

前記制御手段は、共同管理するコンテンツを構成する複数のデータ・ブロックから、自キャッシュ装置における記憶保持対象とする保留データ・ブロックの決定処理を、以下の処理、すなわち、

判定対象データ・ブロックが構成要素となるコンテンツの先頭データ・ブ
25 ックからのオフセット情報としてのブロックオフセット番号：Cに基づいて、

$C \bmod N$ を算出し、

$C \bmod N = 0$ の場合は $V = N$ 、

$C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、

$V = m$ であるか否かを判定し、

V = mである場合に、該データ・ブロックを保留データ・ブロックとして決定する処理、

を実行する構成であることを特徴とする請求項 1 に記載のキャッシュ装置。

5 3. 前記制御手段は、

記憶手段に格納したデータ・ブロックについての管理リストであるLRUリストからの排除対象エントリに対応するデータ・ブロックについて、前記キャッシュグループによる共同管理が可能か否かをキャッシュ装置間のメッセージ送受信により判定し、

10 共同管理可能であると判定した場合に、前記キャッシュグループ・テーブルを生成する構成であることを特徴とする請求項 1 に記載のキャッシュ装置。

4. 前記制御手段は、

15 記憶手段に格納したデータ・ブロックについての管理リストであるLRUリストからの排除対象エントリに対応するデータ・ブロックについて、

前記キャッシュグループ・テーブルの格納情報に基づいて、コンテンツを構成する複数のデータ・ブロック中から、自キャッシュ装置における記憶保持データとしての保留データ・ブロックを決定し、決定した保留データ・ブロック以外のデータ・ブロックについては、削除の許容されるデータ・ブロック管理
20 リストとしてのフリーブロックリストにより管理する構成であることを特徴とする請求項 1 に記載のキャッシュ装置。

5. 前記削除保留リストは、データ・ブロックの保存優先度に応じた複数のリストによって構成され、

25 前記制御手段は、

データ・ブロックの保存優先度を判定し、該判定結果に基づいて選択した削除保留リストに対して前記管理データ・ブロックに対応するエントリの設定を行なう構成であることを特徴とする請求項 1 に記載のキャッシュ装置。

6. 前記制御手段は、

端末からのデータ・ブロック取得要求が、前記キャッシュグループによる共同管理対象のコンテンツに含まれるデータ・ブロック取得要求である場合に、

5 取得要求対象のデータ・ブロックが前記キャッシュグループを構成するいずれのキャッシュ装置の保持対象データ・ブロックであるかを判定し、該判定結果に基づいてデータ・ブロックを自キャッシュ装置または、前記キャッシュグループを構成する他のキャッシュ装置から取得する処理を実行する構成であることを特徴とする請求項 1 に記載のキャッシュ装置。

10 7. 前記キャッシュグループ・テーブルは、キャッシュグループによって共同管理するコンテンツの識別子と、キャッシュグループを構成するキャッシュ装置数：Nと、各キャッシュ装置に対応して設定されるグループ・メンバー番号：mとを有し、

前記制御手段は、

15 取得要求対象のデータ・ブロックが構成要素となるコンテンツの先頭データ・ブロックからのオフセット情報としてのブロックオフセット番号：Cに基づいて、

$C \bmod N$ を算出し、

$C \bmod N = 0$ の場合は $V = N$ 、

20 $C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、

$V = m$ を算出し、

算出したmをグループ・メンバー番号として有するキャッシュ装置からデータ・ブロックの取得処理を実行する構成であることを特徴とする請求項 6 に記載のキャッシュ装置。

25

8. ネットワークを介して受信したデータをキャッシュデータとして格納し、端末からのデータ要求に応じてキャッシュデータを取得して端末に送信するキャッシュ装置におけるキャッシュデータ管理方法であり、

自キャッシュ装置およびネットワークに接続された他のキャッシュ装置に

よって構成され、コンテンツの共同管理を実行するキャッシュグループに関するグループ設定情報を格納したキャッシュグループ・テーブルの格納情報に基づいて、コンテンツを構成する複数のデータ・ブロックから、自キャッシュ装置における記憶保持データとしての保留データ・ブロックを決定する保留データ・ブロック決定ステップと、

決定した保留データ・ブロックの対応情報からなるエントリによって構成される削除保留リストに従って保留データ・ブロックの管理を実行する管理ステップと、

を有することを特徴とするキャッシュデータ管理方法。

10

9. 前記キャッシュグループ・テーブルは、キャッシュグループによって共同管理するコンテンツの識別子と、キャッシュグループを構成するキャッシュ装置数： N と、各キャッシュ装置に対応して設定されるグループ・メンバー番号： m とを有し、

15

前記保留データ・ブロック決定ステップは、共同管理するコンテンツを構成する複数のデータ・ブロックから、自キャッシュ装置における記憶保持対象とする保留データ・ブロックの決定処理を、以下の処理、すなわち、

20

判定対象データ・ブロックが構成要素となるコンテンツの先頭データ・ブロックからのオフセット情報としてのブロックオフセット番号： C に基づいて、

$C \bmod N$ を算出し、

$C \bmod N = 0$ の場合は $V = N$ 、

$C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、

$V = m$ であるか否かを判定し、

25

$V = m$ である場合に、該データ・ブロックを保留データ・ブロックとして決定する処理を実行することを特徴とする請求項8に記載のキャッシュデータ管理方法。

10. 前記キャッシュデータ管理方法は、さらに、

記憶手段に格納したデータ・ブロックについての管理リストであるLRUリストからの排除対象エントリに対応するデータ・ブロックについて、前記キャッシュグループによる共同管理が可能か否かをキャッシュ装置間のメッセージ送受信により判定し、共同管理可能であると判定した場合に、前記キャッシュグループ・テーブルを生成する処理を実行することを特徴とする請求項8に記載のキャッシュデータ管理方法。

11. 前記キャッシュデータ管理方法は、さらに、

10 記憶手段に格納したデータ・ブロックについての管理リストであるLRUリストからの排除対象エントリに対応するデータ・ブロックについて、前記キャッシュグループ・テーブルの格納情報に基づいて、コンテンツを構成する複数のデータ・ブロック中から、自キャッシュ装置における記憶保持データとしての保留データ・ブロックを決定し、決定した保留データ・ブロック以外のデータ・ブロックについては、削除の許容されるデータ・ブロック管理リストとしてのフリーブロックリストにより管理するステップを含むことを特徴とする請求項8に記載のキャッシュデータ管理方法。

12. 前記削除保留リストは、データ・ブロックの保存優先度に応じた複数のリストによって構成され、

20 前記キャッシュデータ管理方法は、さらに、

データ・ブロックの保存優先度を判定し、該判定結果に基づいて選択した削除保留リストに対して前記管理データ・ブロックに対応するエントリの設定を行なうステップを含むことを特徴とする請求項8に記載のキャッシュデータ管理方法。

25

13. 前記キャッシュデータ管理方法は、さらに、

端末からのデータ・ブロック取得要求が、前記キャッシュグループによる共同管理対象のコンテンツに含まれるデータ・ブロック取得要求である場合に、取得要求対象のデータ・ブロックが前記キャッシュグループを構成するいずれ

のキャッシュ装置の保持対象データ・ブロックであるかを判定し、該判定結果に基づいてデータ・ブロックを自キャッシュ装置または、前記キャッシュグループを構成する他のキャッシュ装置から取得するデータ・ブロック取得処理を実行するステップを含むことを特徴とする請求項 8 に記載のキャッシュデータ管理方法。

5

1 4. 前記キャッシュグループ・テーブルは、キャッシュグループによって共同管理するコンテンツの識別子と、キャッシュグループを構成するキャッシュ装置数：Nと、各キャッシュ装置に対応して設定されるグループ・メンバー番号：mとを有し、

10

前記データ・ブロック取得処理は、

取得要求対象のデータ・ブロックが構成要素となるコンテンツの先頭データ・ブロックからのオフセット情報としてのブロックオフセット番号：Cに基づいて、

15

$C \bmod N$ を算出し、

$C \bmod N = 0$ の場合は $V = N$ 、

$C \bmod N \neq 0$ の場合は $V = (C \bmod N)$ として、

$V = m$ を算出し、

算出したmをグループ・メンバー番号として有するキャッシュ装置からデータ・ブロックを取得する処理として実行することを特徴とする請求項 1 3 に記載のキャッシュデータ管理方法。

20

1 5. ネットワークを介して受信したデータをキャッシュデータとして格納し、端末からのデータ要求に応じてキャッシュデータを取得して端末に送信するキャッシュ装置におけるキャッシュデータ管理処理を実行するコンピュータ・プログラムであり、

25

自キャッシュ装置およびネットワークに接続された他のキャッシュ装置によって構成され、コンテンツの共同管理を実行するキャッシュグループに関するグループ設定情報を格納したキャッシュグループ・テーブルの格納情報に基づ

づいて、コンテンツを構成する複数のデータ・ブロックから、自キャッシュ装置における記憶保持データとしての保留データ・ブロックを決定する保留データ・ブロック決定ステップと、

5 決定した保留データ・ブロックの対応情報からなるエントリによって構成される削除保留リストに従って保留データ・ブロックの管理を実行する管理ステップと、

を有することを特徴とするコンピュータ・プログラム。

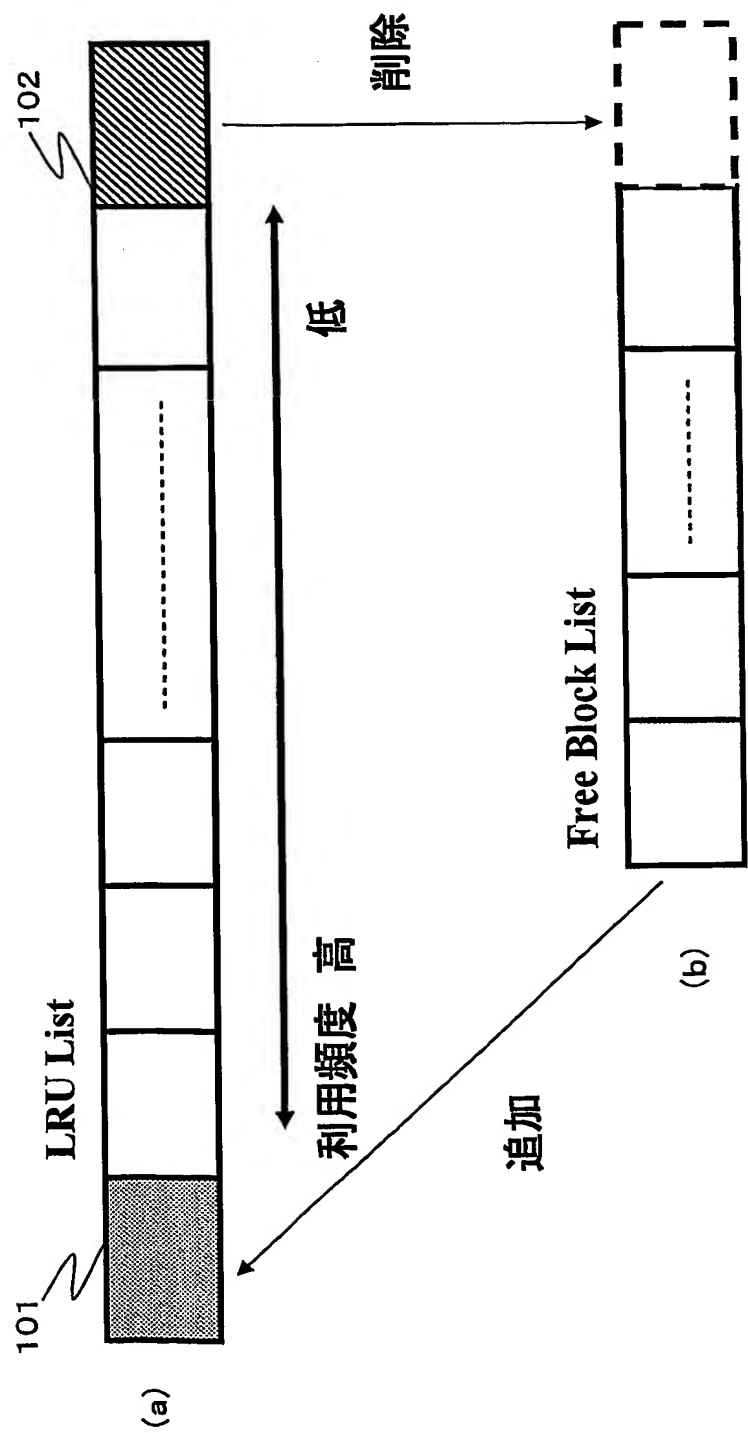
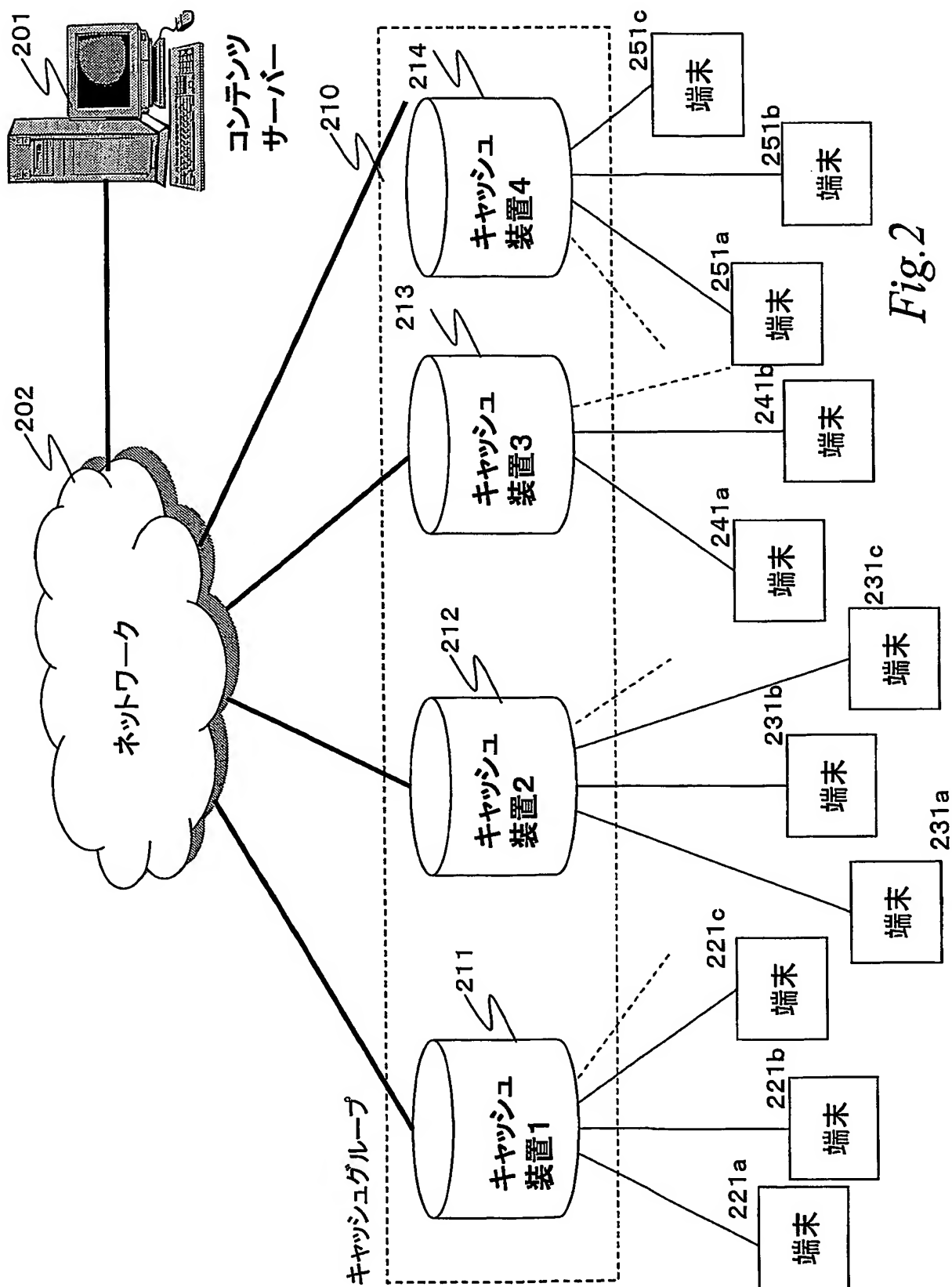
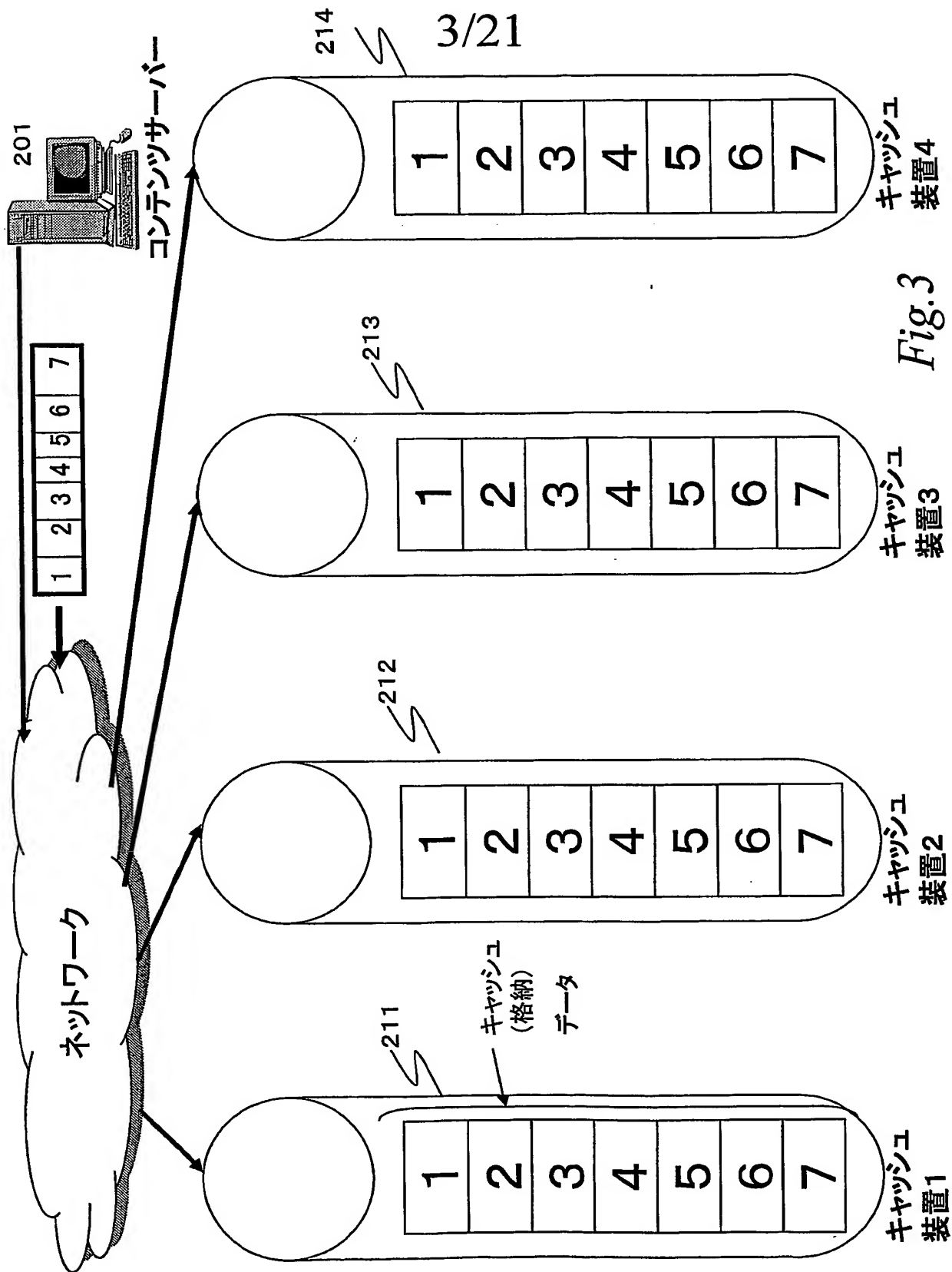


Fig. 1

2/21





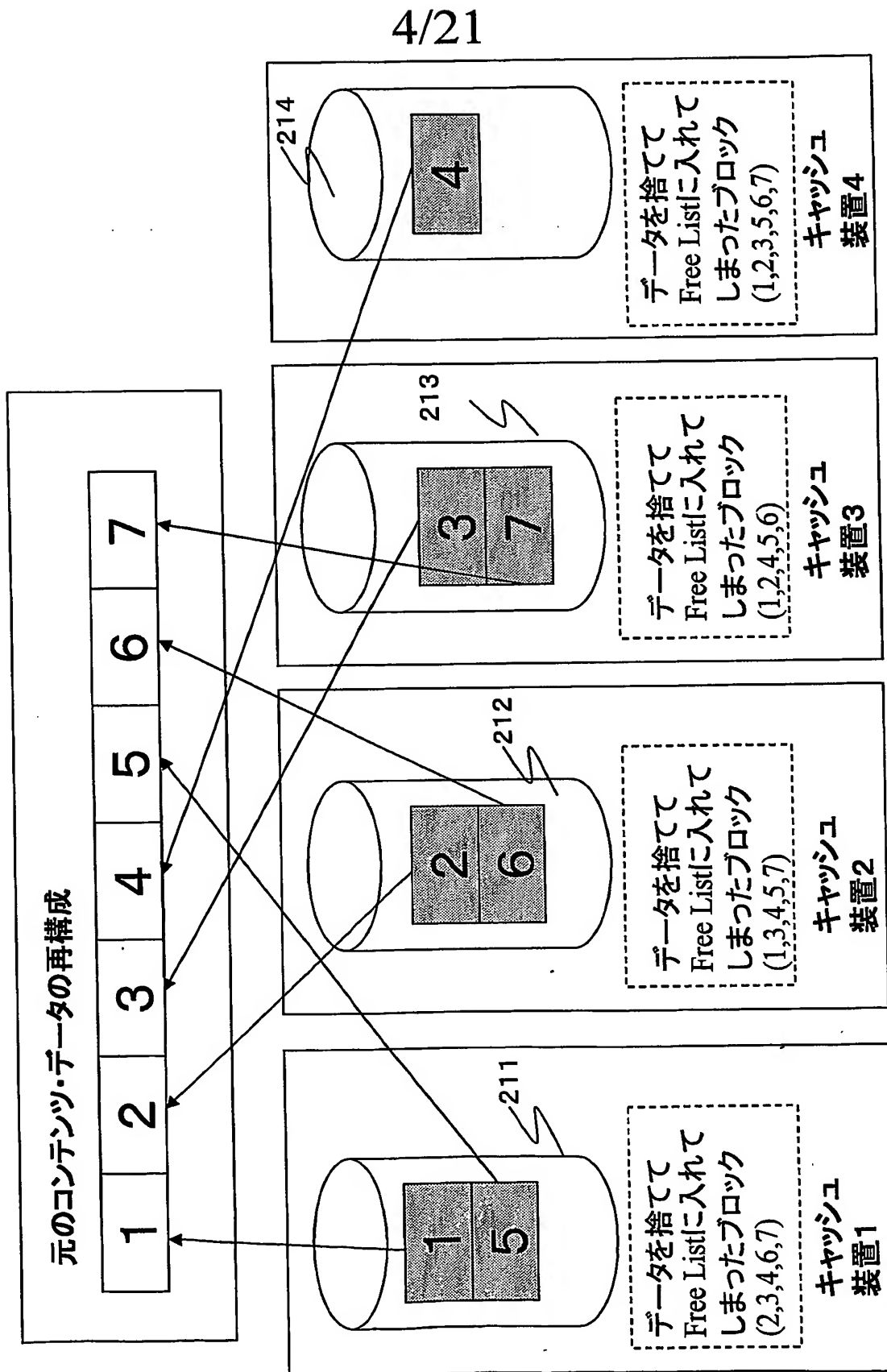


Fig.4

(a)

0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
オペコード(Opcode)										バージョン(Version)										メッセージ長(Message Length)									
リクエストNo. (Request Number)																													
オプション(Options)																													
オプションデータ(Optional Data)																													
送信ホストアドレス(Sender Host Address)																													
ペイロード(Payload)																													

(b)

リクエストホストアドレス(Request Host Address)
ヌル・ターミネートURL(Null-Terminated URL)

Fig.5

6/21

ホスト名	IP Address	ICP ポート番号
cache1.home1.com	xx.yy.zz.a1	3130
cache2.home2.com	xx.yy.zz.a2	3130
cache3.home3.com	xx.yy.zz.a3	3130
.....

Fig.6

7/21

0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
リクエストホストアドレス(Request Host Address)																					
ヌルターミネート(Null-Terminated)コンテンツ識別子																					
メンバー数 (N)			メンバー番号 (m)					グループ ID													
キャッシュメンバーバ(Cache Member) #1 IPアドレス (IP Address)																					
キャッシュメンバーバ(Cache Member) #2 IPアドレス (IP Address)																					
キャッシュメンバーバ(Cache Member) #3 IPアドレス (IP Address)																					
キャッシュメンバーバ(Cache Member) #4 IPアドレス (IP Address)																					
.....																					
キャッシュメンバーバ(Cache Member) #N IPアドレス (IP Address)																					

Fig. 7

8/21

対象とする コンテント・データ の識別子	グループ 構成メンバー数(N)	グループ ID	グループメンバーリスト	
			メンバー番号(m)	IPアドレス
12345678	4	231	1	xx.yy.zz.a1
			2	xx.yy.zz.a2
			3	xx.yy.zz.a3
			4	xx.yy.zz.a4
34567890	3	232	1	xx.yy.zz.b1
			2	xx.yy.zz.b2
			3	xx.yy.zz.b3
...

Fig.8

9/21

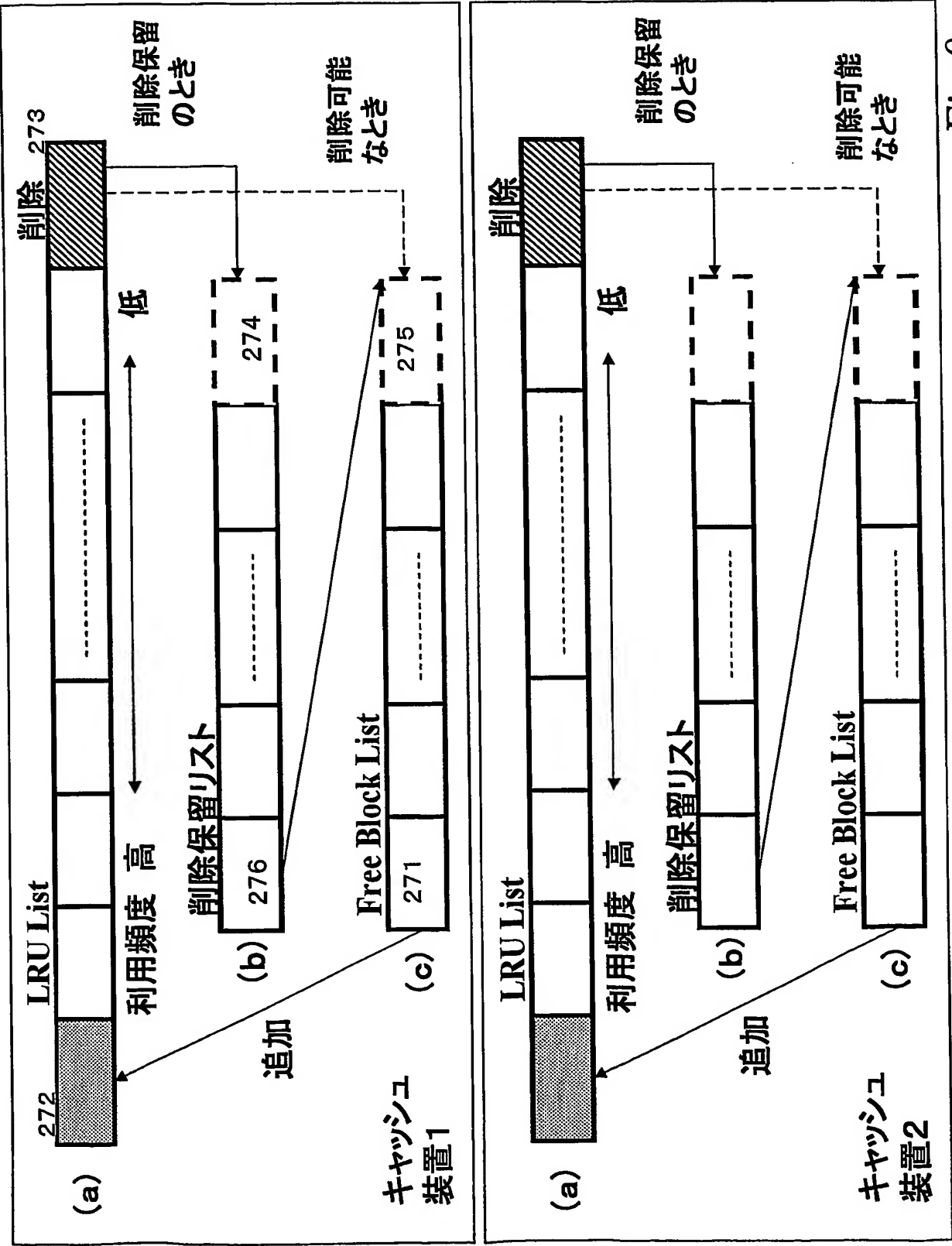
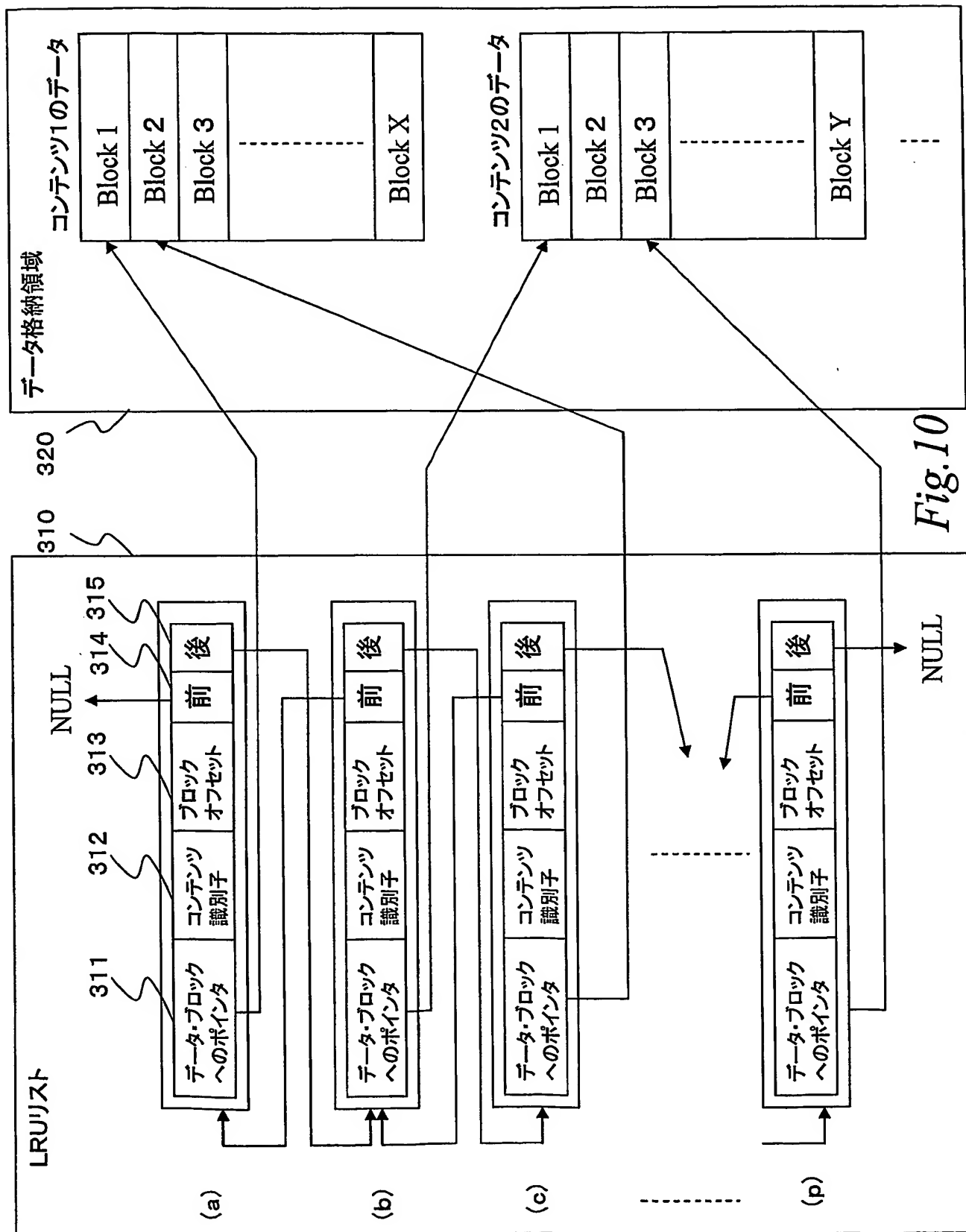
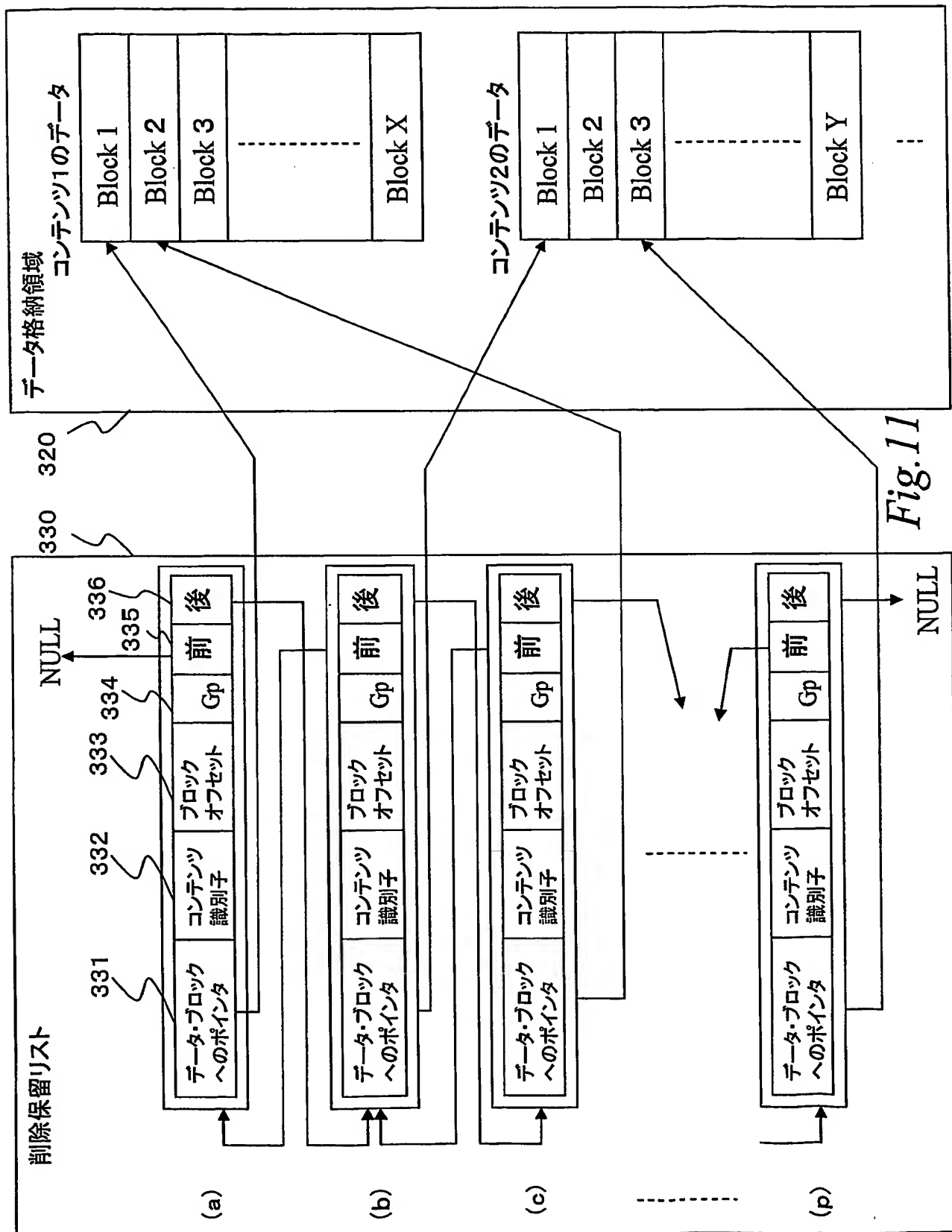


Fig.9

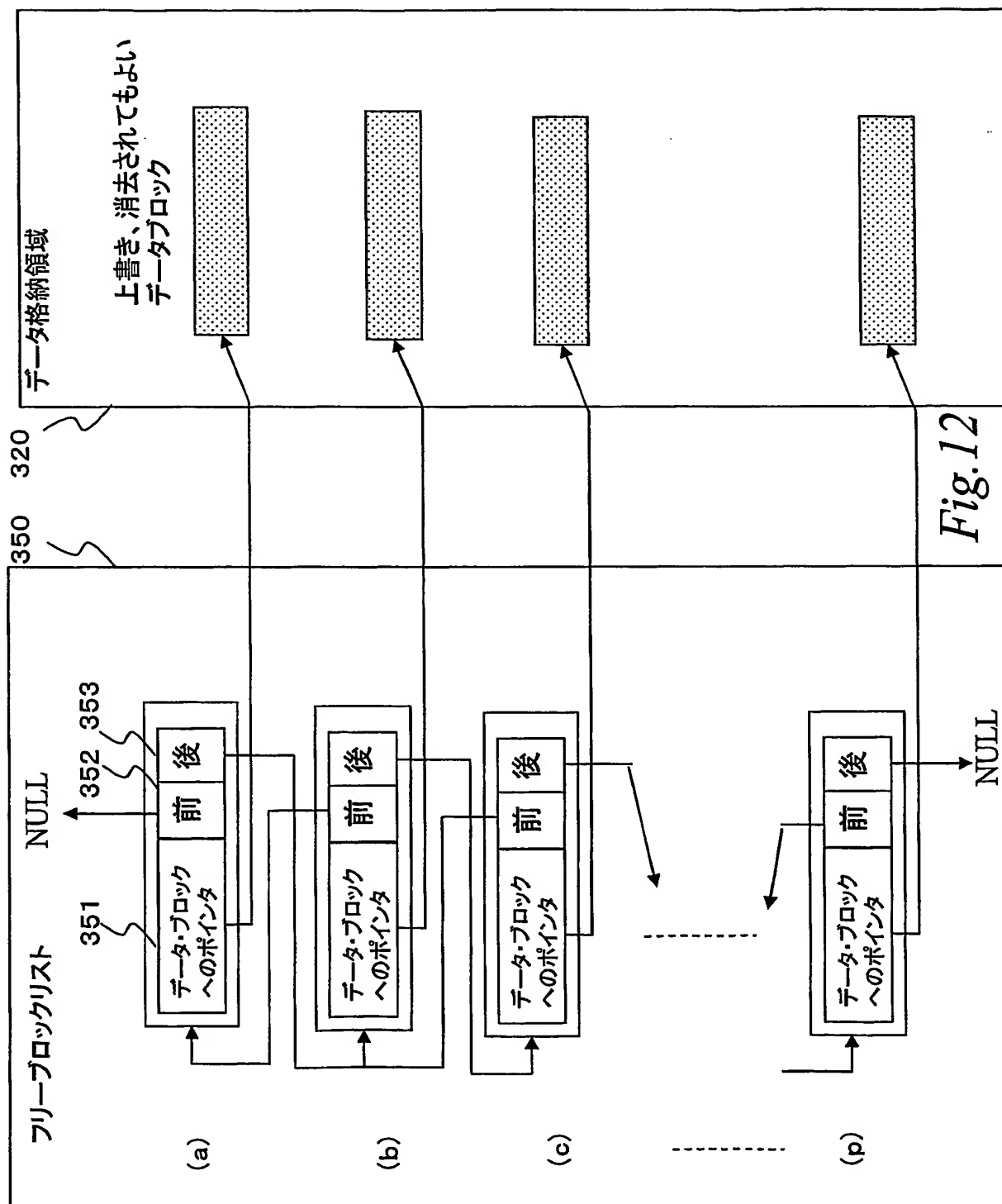
10/21



11/21



12/21



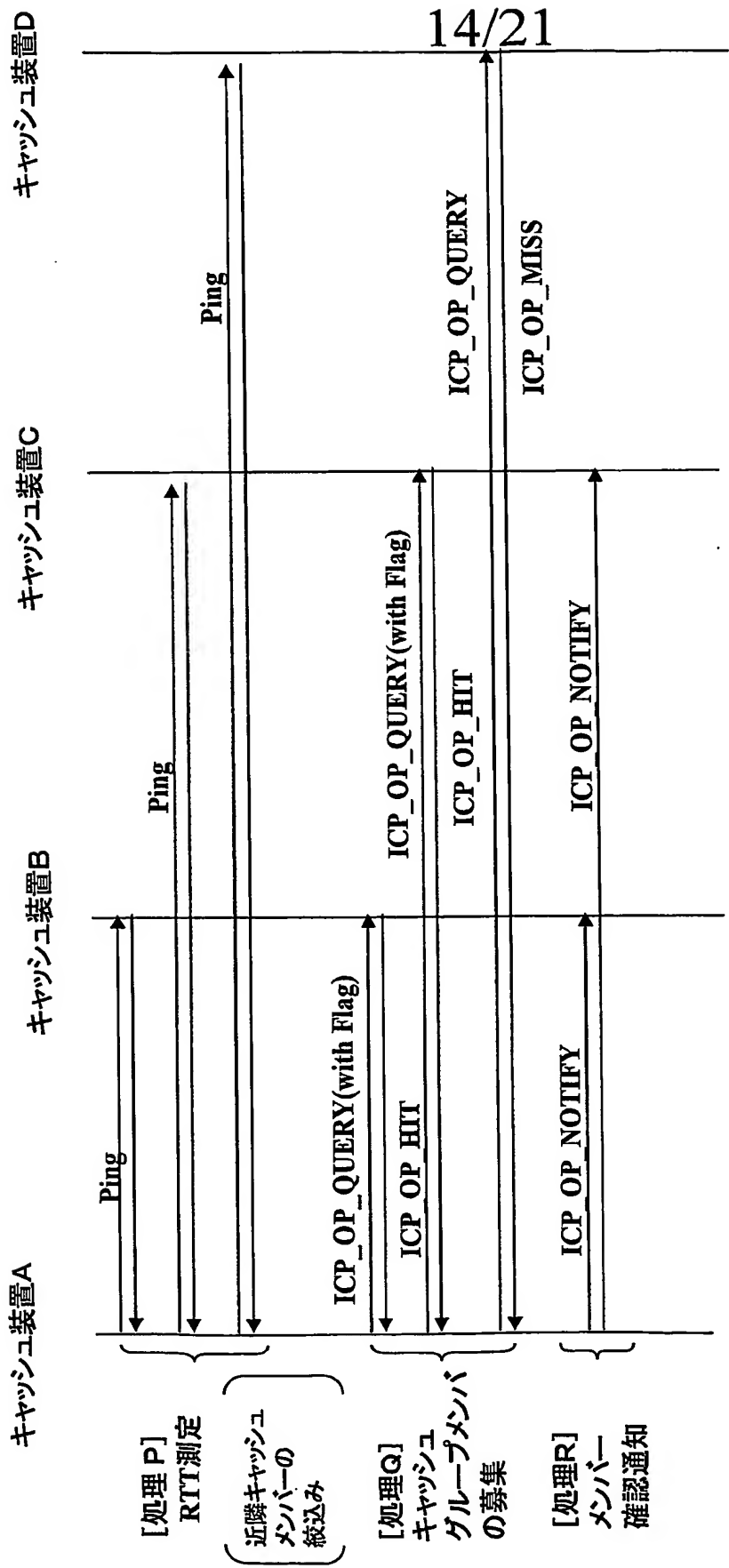
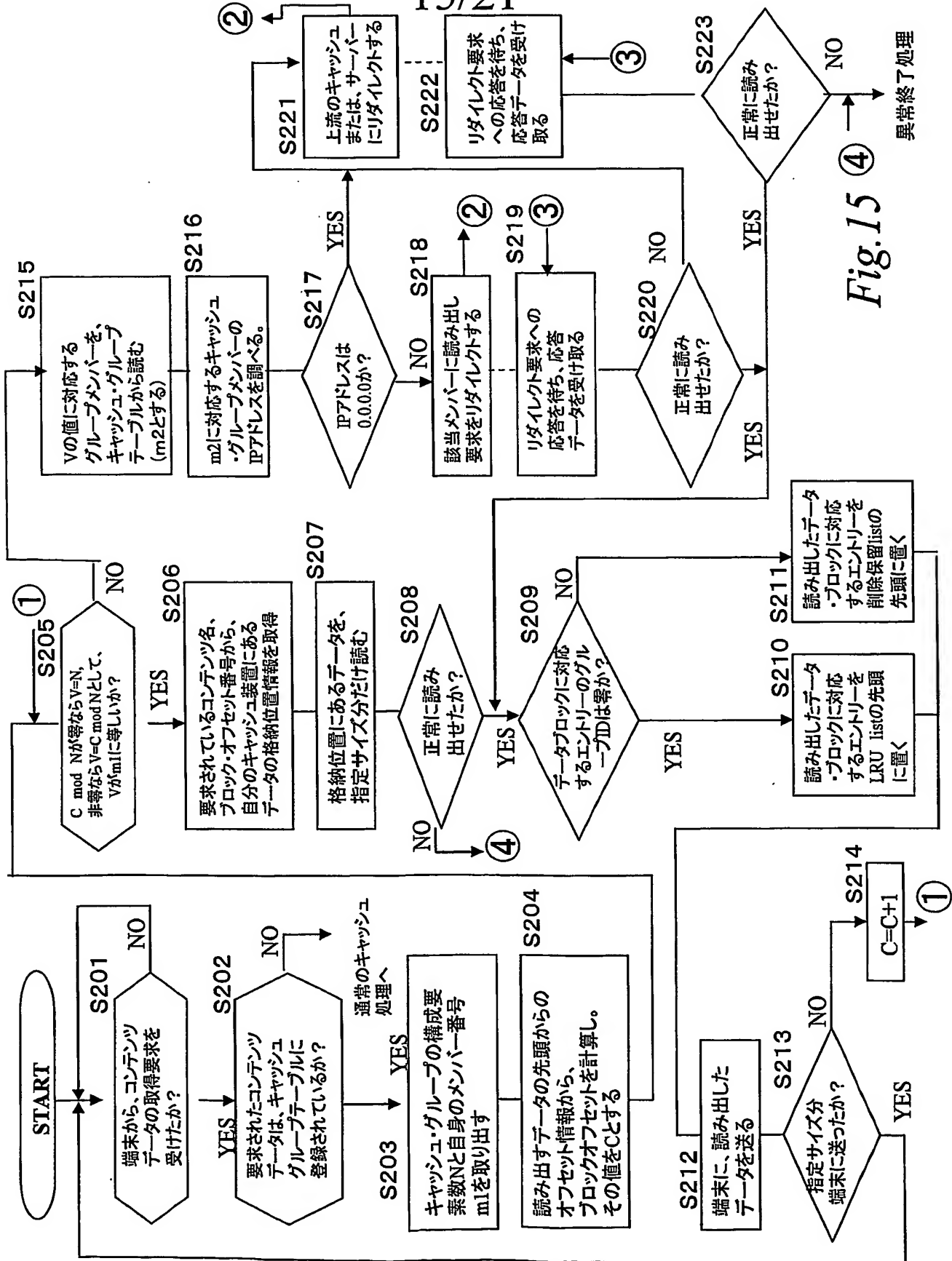


Fig.14

15/21



16/21

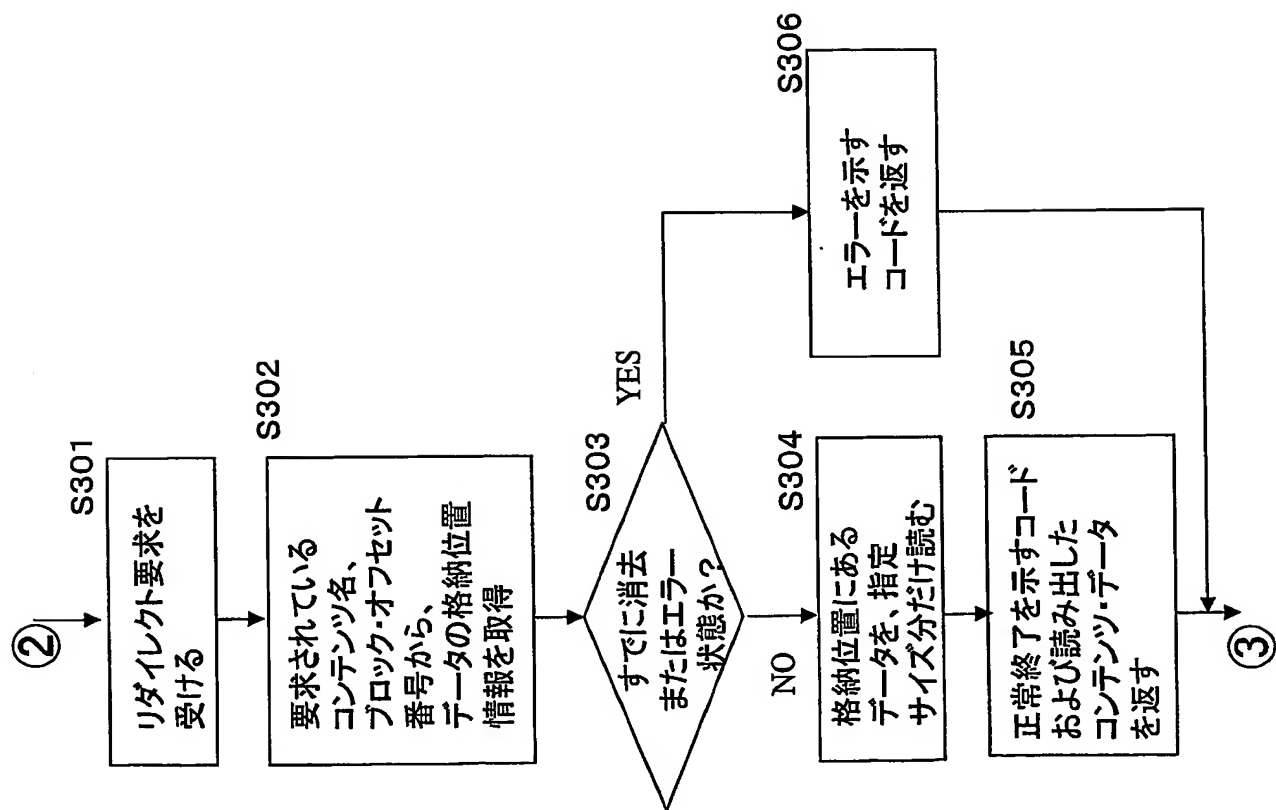


Fig. 16

上流キャッシュ装置
または、サーバー

グループ・メンバー
のキャッシュ装置2

グループ・メンバー
のキャッシュ装置1

アクセス要求を出した端末
が接続しているキャッシュ装置A

端末

[処理S]
コンテンツ
データの
取得
(ヒット時)

[処理T]
コンテンツ
データの
取得
(ミス時)

[処理U]
グループ
解消通知

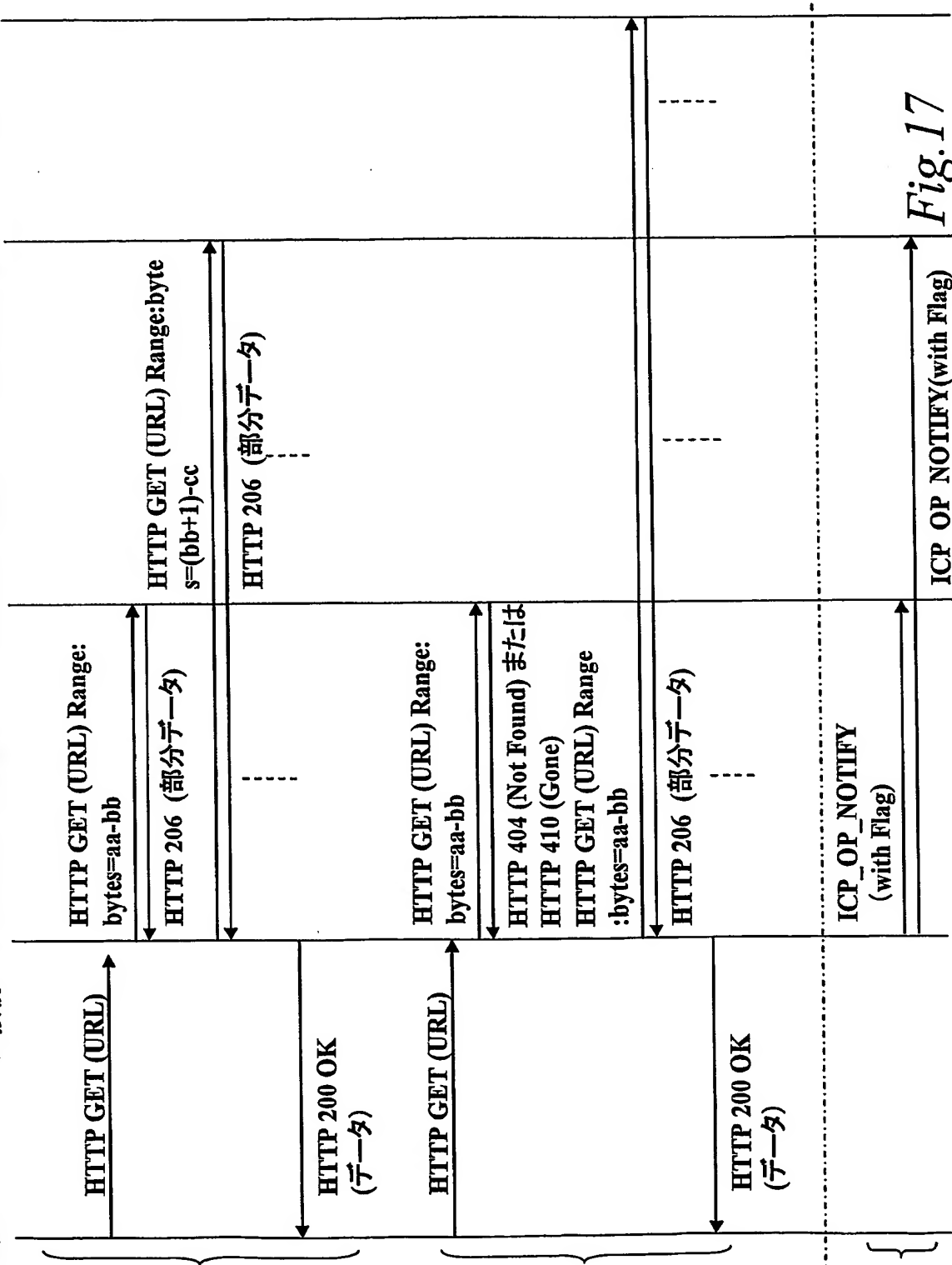


Fig.17

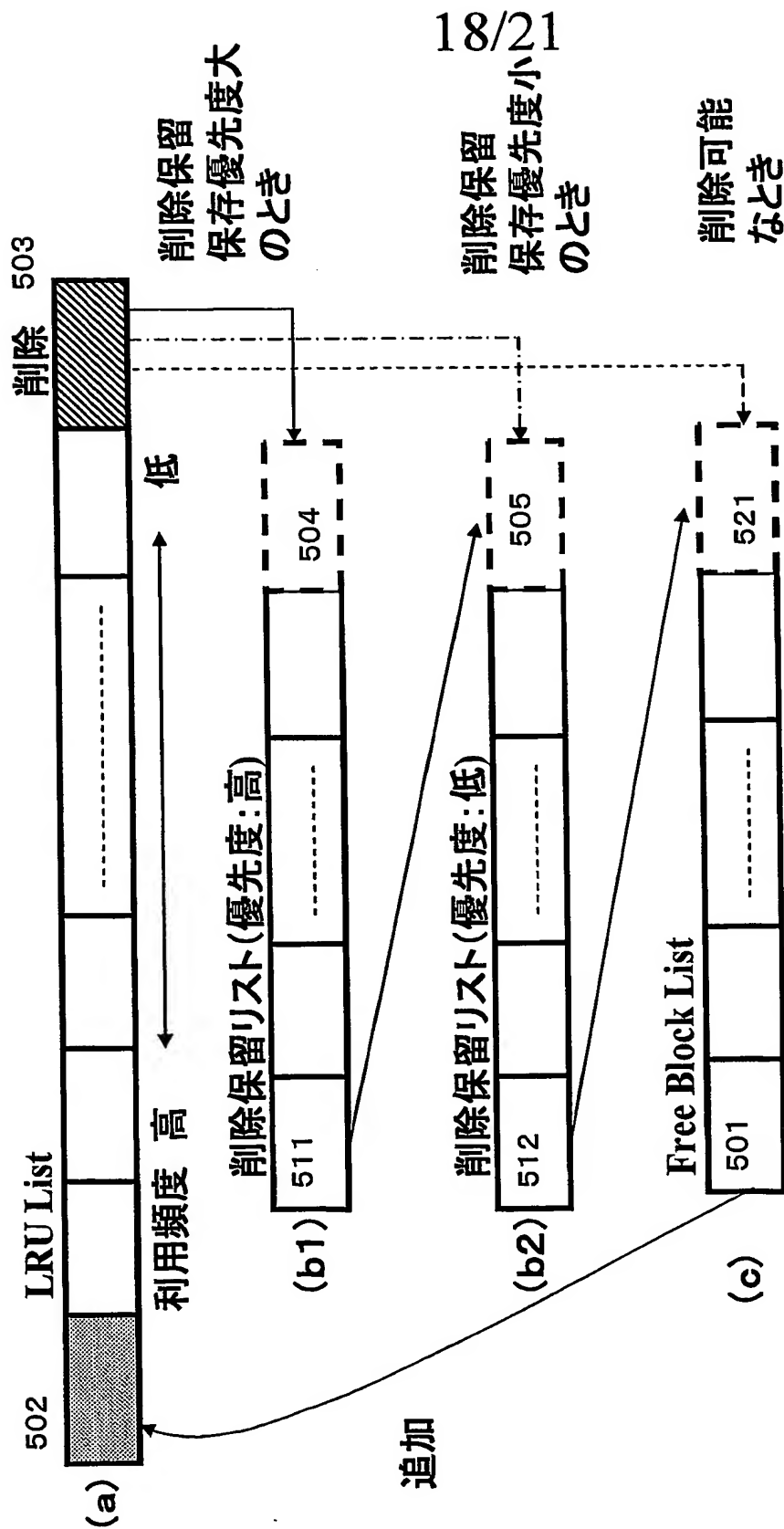
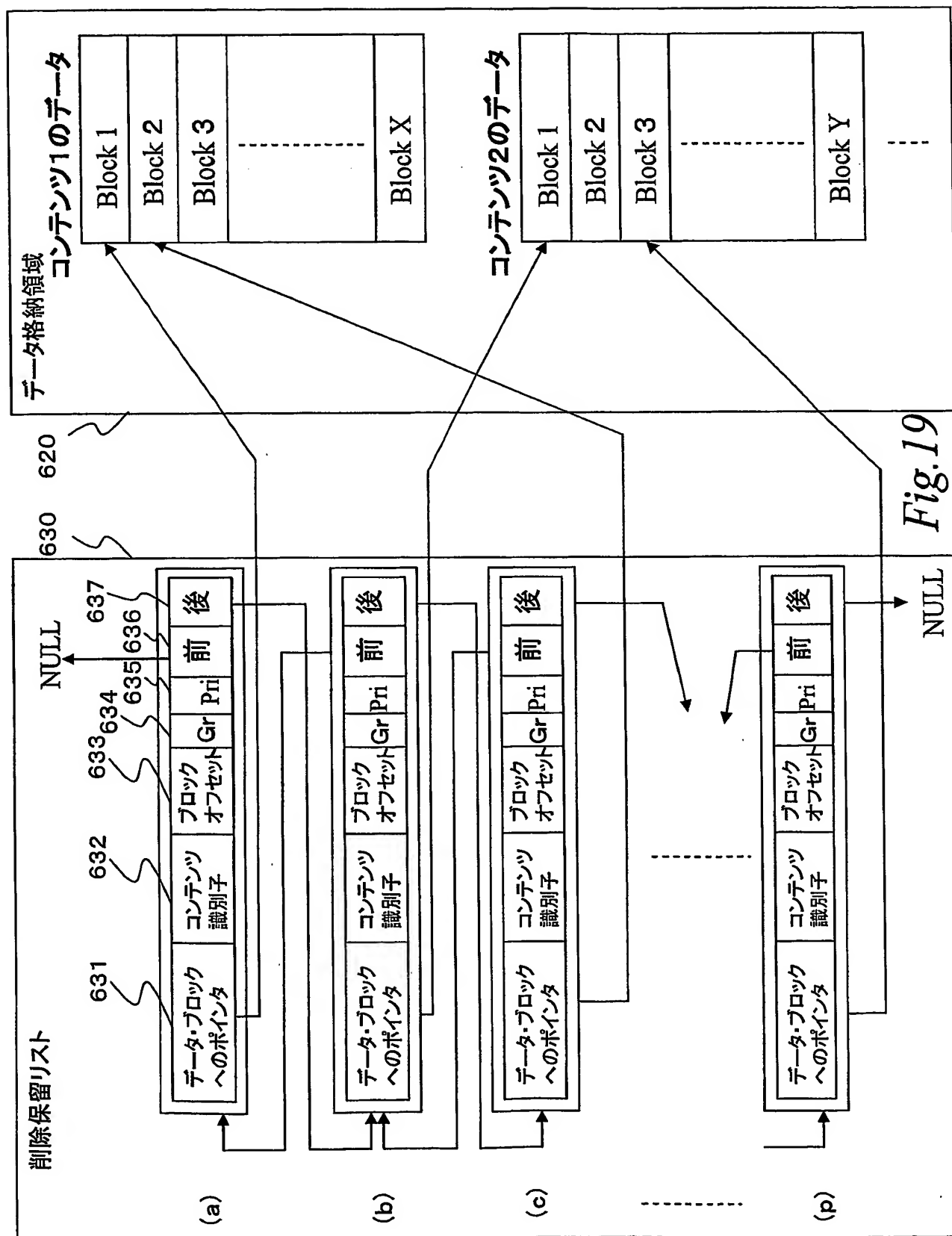


Fig.18

19/21



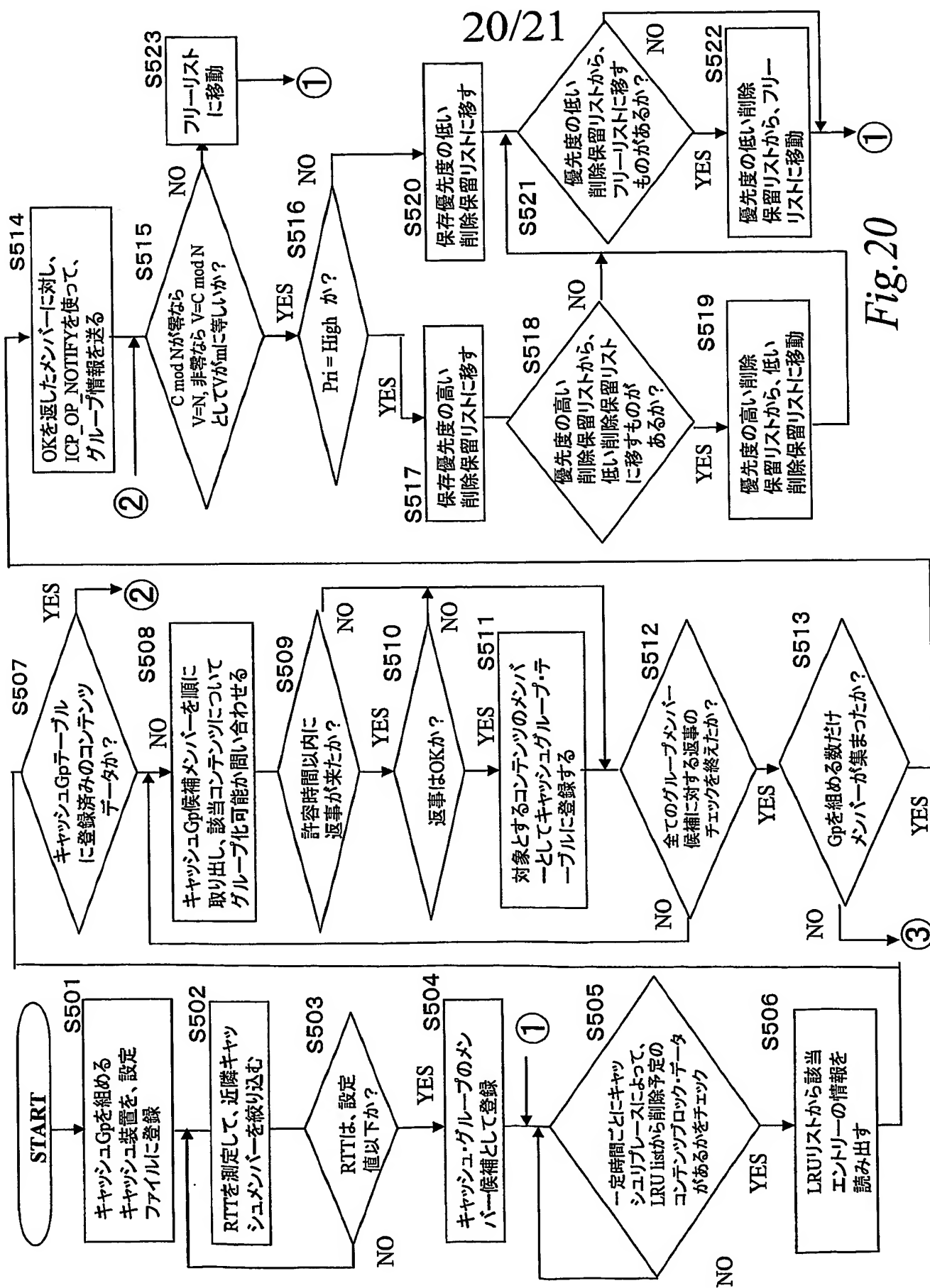


Fig. 20

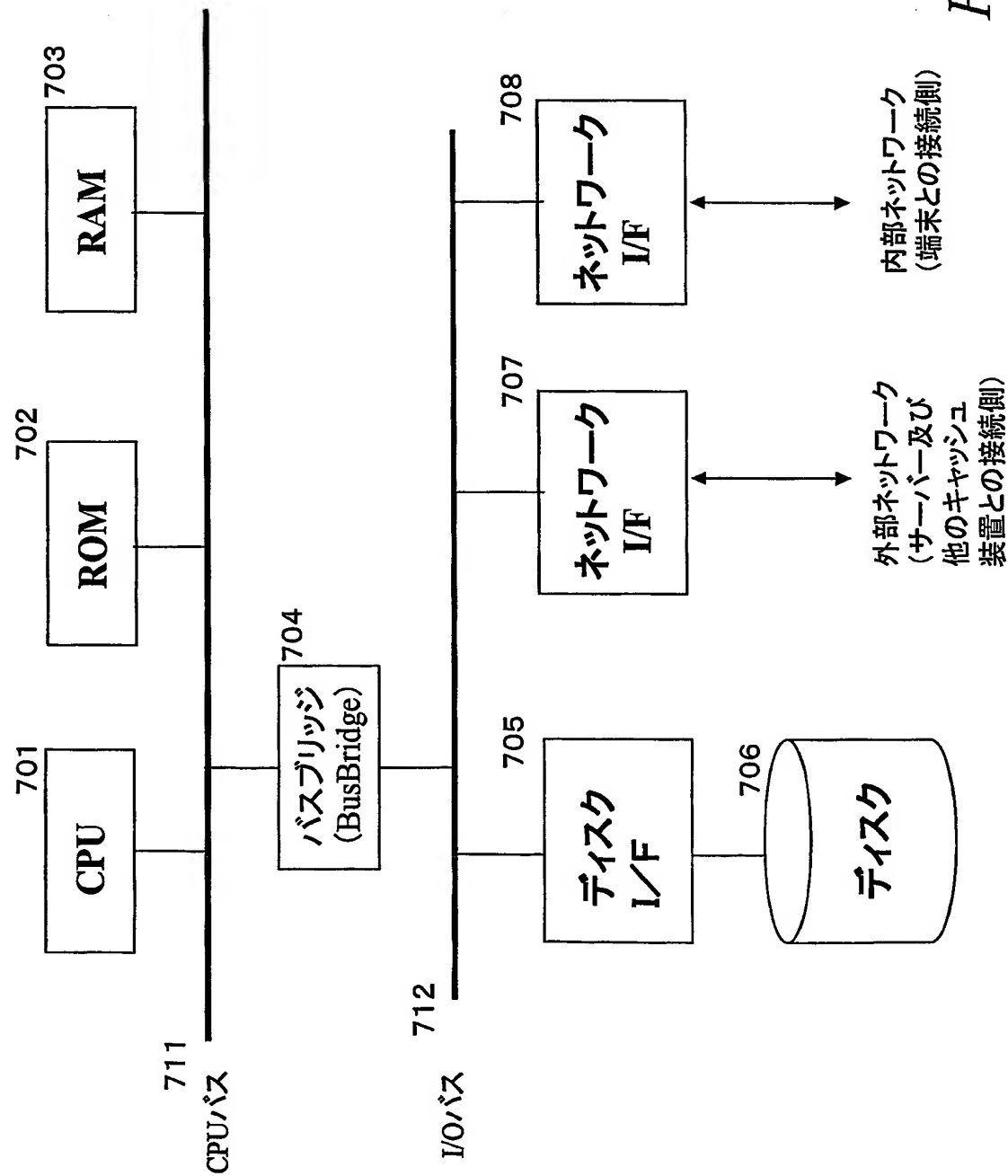


Fig. 21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP03/07971

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G06F12/00, 13/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl⁷ G06F12/00, 13/00, 17/30

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2003
Kokai Jitsuyo Shinan Koho	1971-2003	Jitsuyo Shinan Toroku Koho	1996-2003

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 10-222411 A (Nippon Telegraph And Telephone Corp.), 21 August, 1998 (21.08.98), Full text; all drawings (Family: none)	1-15
A	JP 2001-202330 A (Matsushita Electric Industrial Co., Ltd.), 27 July, 2001 (27.07.01), Full text; all drawings & EP 1107533 A2	1-15
A	JP 2002-149477 A (Sony Corp.), 24 May, 2002 (24.05.02), Full text; all drawings (Family: none)	1-15

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:
 "A" document defining the general state of the art which is not considered to be of particular relevance
 "E" earlier document but published on or after the international filing date
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 "O" document referring to an oral disclosure, use, exhibition or other means
 "P" document published prior to the international filing date but later than the priority date claimed

"I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
 "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
 "&" document member of the same patent family

Date of the actual completion of the international search
18 August, 2003 (18.08.03)

Date of mailing of the international search report
02 September, 2003 (02.09.03)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC))
Int.Cl.⁷ G 0 6 F 1 2 / 0 0, 1 3 / 0 0

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))
Int.Cl.⁷ G 0 6 F 1 2 / 0 0, 1 3 / 0 0, 1 7 / 3 0

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年
日本国公開実用新案公報 1971-2003年
日本国登録実用新案公報 1994-2003年
日本国実用新案登録公報 1996-2003年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 10-222411 A(日本電信電話株式会社) 1998. 08. 21, 全文, 全図(ファミリーなし)	1-15
A	JP 2001-202330 A(松下電器産業株式会社) 2001. 07. 27, 全文, 全図 & EP 1107533 A2	1-15
A	JP 2002-149477 A(ソニー株式会社) 2002. 05. 24, 全文, 全図(ファミリーなし)	1-15

☐ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
「O」 口頭による開示、使用、展示等に言及する文献
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」 同一パテントファミリー文献

国際調査を完了した日 18. 08. 03

国際調査報告の発送日 02.09.03

国際調査機関の名称及びあて先
日本国特許庁 (ISA/J P)
郵便番号 100-8915
東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)
相崎 裕恒

5 N 9 2 9 0

電話番号 03-3581-1101 内線 3545

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.